

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

**Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу**

До захисту допущено:

В.о.завідувача кафедри

_____ Оксана ТИМОЩУК

«__» _____ 20__ р.

Дипломна робота
на здобуття ступеня бакалавра
за освітньо-професійною програмою «Системний аналіз і управління»
спеціальності 124 «Системний аналіз»
на тему: «Моделі та прогнози процесу поширення коронавірусу»

Виконала:

студентка IV курсу, групи КА-64

Мельничук Анастасія Володимирівна _____

Керівник:

професор, д.т.н.

Бідюк П.І. _____

Консультант з економічного розділу:

доцент, к.е.н.

Шевчук О.А. _____

Рецензент:

Професор, д.т.н.

Теленик Сергій Федорович _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студентка _____

Київ – 2020 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут прикладного системного аналізу
Кафедра математичних методів системного аналізу

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 124 «Системний аналіз»

Освітньо-професійна програма «Системний аналіз і управління»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ О.Л. Тимощук

«___» _____ 20__ р.

ЗАВДАННЯ

на дипломну роботу студенту

Мельничук Анастасія Володимирівна

1. Тема роботи «Моделі та прогнози процесу поширення коронавірусу», керівник роботи Бідюк Петро Іванович, професор, д.т.н., затверджені наказом по університету від «___» _____ 20__ р. № _____
2. Термін подання студентом роботи _____
3. Вихідні дані до роботи
4. Зміст роботи
5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)
6. Консультанти розділів роботи*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Економічний	Шевчук О.А., доцент		

* Якщо визначені консультанти. Консультантом не може бути зазначено керівника дипломної роботи.

7. Дата видачі завдання _____

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Формулювання тематики (напрямку) дослідження.	03.09.2019 – 30.09.2019	
2	Аналіз актуальності задач стосовно тематики дослідження	01.10.2019 – 30.10.2019	
3	Аналіз відомих результатів стосовно тематики дослідження	01.11.2019 – 30.11.2019	
4	Формулювання задач дослідження	01.12.2019 – 30.12.2019	
5	Уточнення теми дипломної роботи	25.02.2019	
6	Збір статичних даних, попередній аналіз даних	01.03.2020 – 30.03.2020	
7	Розробка програмного продукту для виконання обчислювальних експериментів	01.03.2020 – 30.04.2020	
8	Виконання обчислювальних експериментів, аналіз та оформлення результатів	01.05.2020 – 20.05.2020	
9	Оформлення пояснювальної записки у цілому	21.05.2020 – 31.05.2020	
10	Підготовка презентації для захисту	28.05.2020 – 01.06.2020	
11	Попередній захист дипломної роботи	01.06.2020 – 03.06.2020	
12	Захист дипломної роботи	15.06.2020 – 18.06.2020	

Студентка

А.В. Мельничук

Керівник

П.І. Бідюк

РЕФЕРАТ

Дипломна робота: 165 с., 40 рис., 29 табл., 14 джерел, 2 додатки

АР, АРУГ, АВТОРЕГРЕСІЯ З ТРЕНДОМ, МНК, СКП, САПП, КОЕФІЦІЄНТ ТЕЙЛА, НЕЛІНІЙНІ НЕСТАЦІОНАРНІ ПРОЦЕСИ, КРИТЕРІЇ АДЕКВАТНОСТІ, ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ СИМУЛЯЦІЙ ПРОЦЕСІВ.

Тема: Моделі та прогнози поширення процесів поширення коронавірусу.

Об'єкт дослідження: процеси розвитку та поширення епідемій, які були спричинені вірусними захворюваннями.

Предмет дослідження: регресійний аналіз статистичних даних, імітаційне моделювання процесу поширення вірусу в різних умовах.

Мета роботи: оцінка наявних статистичних даних та побудова регресійних моделей для даних часових рядів; побудова короткочасного однокрокового прогнозу кумулятивної кількості тих, хто інфікований і тих, хто одужав; оцінка та аналіз основних методів боротьби проти поширення епідемій вірусних захворювань.

Методи дослідження: методи статистичного та регресійного аналізу даних (часових рядів), модулі для побудови симуляцій в імітаційному моделюванні.

В роботі розглянуто проблему оцінки та аналізу процесів поширення вірусів, на основі підібраних математичних моделей для них побудовано короткостроковий прогноз, проведено порівняння методів боротьби з епідеміями за допомогою створення імітаційних моделей примітивного варіанту дії.

Для аналізу і прогнозування використані реальні дані, взяті з репозиторію даних Центру системної науки і техніки імені Джонса Хопкінса.

Всі обчислення виконані за допомогою пакету для статистичної обробки даних Eviews та модулів для побудови симуляцій в AnyLogic. Параметри моделей оцінені за допомогою методу найменших квадратів(МНК). Для аналізу адекватності моделей використовувалися такі критерії якості оцінювання як: коефіцієнт детермінації, сума квадратів похибок, статистика Дарбіна-Уотсона. Для аналізу якості отриманих прогнозів використовувалися такі критерії як: середньоквадратична похибка(СКП), середня абсолютна похибка у відсотках (САПП) та коефіцієнт Тейла.

На основі отриманих даних були побудовані графіки для візуального аналізу та оцінки процесу поширення коронавірусу. Здійснено порівняльний аналіз результатів, виконаних за допомогою різних інструментаріїв.

ABSTRACT

Thesis: 165 p., 40 fig., 29 tables, 14 sources, 2 appendixes.

AR, ARMA, AUTOREGRESSIONS WITH TREND, LSM, SSE, MAPE, COEFFICIENT TAYLE, NONLINEAR PROCESSES, ADEQUACY CRITERIA, SIMULATION MODELLING, SIR, VIR.

The theme: Models and forecasts of coronavirus propagation processes.

Object of research: processes of development and spread of epidemics caused by viral diseases.

Subject of research: regression analysis of statistical data, simulation modeling of virus propagation process in different conditions.

Purpose of work: evaluation of available statistical data and construction of regression models for time series data; construction of short-term one-step prognosis of cumulative number of those who are infected and who have recovered; evaluation and analysis of the main methods to control the spread of viral disease epidemics.

Methods of research: methods of statistical and regression data analysis (time series), modules for building simulations in simulation modeling.

The problem of estimation and analysis of processes of virus spreading is considered in the work, on the basis of selected mathematical models the short-term forecast is constructed for them, comparison of methods of struggle against epidemics by means of creation of imitation models of a primitive variant of action is carried out.

Real data taken from the Johns Hopkins Center for Systems Science and Technology data repository is used for analysis and forecasting.

All calculations were performed using the EvIEWS statistical processing package and modules for building simulations in AnyLogic. The model parameters were estimated using the least squares method (LSM). To analyze the adequacy of the models, the following evaluation quality criteria were used: determination coefficient, sum of error squares, Darbin-Watson statistics. The following criteria

were used for the quality analysis of the obtained forecasts: mean square error (MSE), mean absolute percentage error (MAPE) and the Tayle coefficient.

Based on the obtained data, graphs were constructed for visual analysis and assessment of the coronavirus distribution process. A comparative analysis of the results performed using various tools.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ.....	10
ПОСТАНОВКА ЗАДАЧІ	11
ВСТУП.....	12
 РОЗДІЛ 1: ОСОБЛИВОСТІ РОЗВИТКУ ПРОЦЕСІВ	
РОЗПОВСЮДЖЕННЯ ПАНДЕМІЙ ТА ЇХ МОДЕЛЮВАННЯ.....	15
1.1 Актуальність проблеми, характеристика розвитку процесів поширення пандемій в Україні та світі.....	15
1.2 Історія виникнення та поширення найбільших пандемій у світі до 20 ст.	20
1.3 Історія виникнення та поширення найбільших пандемій у світі 20-21 ст.	27
1.4 Принципи моделювання епідемії	32
ВИСНОВКИ ДО РОЗДІЛУ	36
 РОЗДІЛ 2 :ВИБІР І ОПИС СТРУКТУР МАТЕМАТИЧНИХ МОДЕЛЕЙ	
ПРОЦЕСІВ ПОШИРЕННЯ ПАНДЕМІЇ.....	39
2.1 Підходи до математичного моделювання вибраних процесів	39
2.2 Методика побудови моделей вибраних процесів	41
2.2.1 Узагальнена класифікація динамічних процесів	41
2.2.2 Методика моделювання нелінійних нестационарних процесів	42
2.2.3 Робота з невизначеностями	47
2.2.4 Модель авторегресії (AR).....	54
2.2.5 Авторегресія з ковзним середнім порядком (p, q) (АРКС)	55
2.2.6 Загальний розв'язок SIR-модель	57
2.3 Інструменти для моделювання і прогнозування обраних процесів	59
 РОЗДІЛ 3: ПОБУДОВА МОДЕЛЕЙ ТА ОЦІНЮВАННЯ ПРОГНОЗІВ	
ПРОЦЕСІВ ПОШИРЕННЯ ПАНДЕМІЇ.....	63
3.1 Загальний підхід до моделювання ситуації поширення пандемії вірусу	63
3.2 Побудова математичних моделей процесів поширення коронавірусу	71
3.2.1 Побудова моделей для прогнозування числа тих, хто захворів в Україні	71
3.2.2 Побудова моделей для прогнозування числа тих, хто одужав в Україні	83
3.2.3 Побудова моделей для прогнозування числа тих, хто захворів в Китаї	91
3.2.4 Побудова моделей для прогнозування числа тих, хто одужав в Китаї	102
3.2.5 Побудова моделей для прогнозування числа тих, хто захворів в Іспанії	104
3.2.6 Побудова моделей для прогнозування числа тих, хто одужав в Іспанії	107

3.2.7 Побудова моделей для прогнозування числа тих, хто захворів в Бразилії	109
3.2.8 Побудова моделей для прогнозування числа тих, хто одужав в Бразилії	112
3.2.9 Побудова моделей для прогнозування числа тих, хто захворів в США	115
3.2.10 Побудова моделей для прогнозування числа тих, хто одужав в США	118
3.3 Альтернативна модель	120
3.4 Моделювання процесу поширення пандемії в програмі AnyLogic	122
3.4.1 Моделювання базової симуляції	123
3.4.2 Моделювання симуляції з центральною точкою відвідування	129
ВИСНОВКИ ДО РОЗДІЛУ	133
РОЗДІЛ 4: ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ	
ПРОГРАМНОГО ПРОДУКТУ	135
4.1 Постановка задачі проектування програмного продукту для моделювання та прогнозу поширення коронавірусу	135
4.2 Обґрунтування функцій програмного продукту	135
4.3 Економічний аналіз варіантів розробки	141
4.4. Вибір кращого варіанта ПП техніко-економічного рівня	146
ВИСНОВКИ ДО РОЗДІЛУ	146
ВИСНОВКИ ДО РОБОТИ	148
ЛІТЕРАТУРА	151
ДОДАТОК А (Статистичні дані)	153
ДОДАТОК В (Презентація)	160

ПЕРЕЛІК СКОРОЧЕНЬ

АР – авторегресія (AR - autoregression)

АРКС – авторегресія ковзного середнього (ARMA – autoregressive with moving average)

МНК – метод найменших квадратів (LSM – least squares method)

РР – різницеві рівняння

САП – середня абсолютна похибка (MAE – mean absolute error)

САПП – середня абсолютна похибка у відсотках (MAPE – mean absolute percent error)

СКП – середньоквадратична похибка (RMSE – root mean squared error)

DW – (Darbin-Watson statistics) статистика Дарбіна-Уотсона

ПОСТАНОВКА ЗАДАЧІ

1. Знайти і підготувати статистичні дані стосовно кількості інфікованих та тих, хто вже одужав в Україні, Китаї, Іспанії, Бразилії, США. За необхідності виконати їх нормування чи логарифмування.
2. Попередньо проаналізувати знайдені часові ряди на наявність тренду.
3. Побудувати апроксимацію процесів поширення коронавірусу за допомогою моделей авторегресії та авторегресії з ковзним середнім.
4. Виконати оцінку короткострокових прогнозів процесів поширення коронавірусу.
5. Проаналізувати отримані результати, продемонструвавши їх на графіках.
6. Провести загальний аналіз умов розвитку та поширення інфекційного захворювання в популяції (в суспільстві).
7. Визначити основні фактори, які впливатимуть на швидкість розповсюдження вірусу та масштаби ураження, на основі яких сформулювати гіпотези.
8. Побудувати різного роду симуляції за допомогою модулів, на основі SIR-моделі, в програмі AnyLogic для підтвердження або спростування висунутих гіпотез та оцінювання ефективності основних методів боротьби проти поширення інфекційних захворювань.
9. Продемонструвати наглядно отримані результати, провести їх аналіз.

ВСТУП

Спалах епідемії, викликаний коронавірусом SARS-CoV-2, представляє великий інтерес для дослідників в зв'язку з високими темпами поширення інфекції і значним числом смертельних випадків. Детальний науковий аналіз цього феномена ще належить провести, але громадськість вже цікавлять питання тривалості епідемії, очікувана кількість хворих і смертей. Для тривалих прогнозів необхідні складні математичні моделі, які вимагають багатьох зусиль для ідентифікації та обчислення невідомих параметрів.

Взагалі епідеміологічні прогнози для вірусологів будують на різній тривалості періоди, виходячи з того, які початкові цілі ставлять дослідники. Наприклад, короткостроковий прогноз (не більше декількох тижнів) застосовують при виявленні епідеміологічного спалаху інфекції для використання отриманих результатів в прийнятті швидких рішень щодо нормалізації ситуації або хоча б мінімізації потенційних масштабів ураження.

Прогноз терміном від двох місяців до півроку називають середньостроковим і рахується найбільш оптимальним та корисним своїми результатами для ведення тактичної боротьби проти поширення епідемії інфекційних захворювань. Такий прогноз поступається трохи точністю короткостроковому, але дозволяє за достатній для підготовки термін передбачити можливі надзвичайні ситуації та прийняти превентивні міри. Найменш якісними вважають довгострокові прогнози тому, що досягти для них високої точності передбачення майже неможливо. Та в стратегічному плануванні без прогнозу на рік чи декілька вперед складно було б оцінити об'єми виробництва лікарських засобів, вакцин, встановлення обладнання в медичних закладів і ресурсів для підготовки медичного персоналу в необхідних кількостях.

Якість прийнятих рішень прямо пропорційне якості прогнозування поведінки їх в майбутньому. Основною метою моделювання процесів

поширення вірусів є якраз їх подальше прогнозування, тобто інформація стосовно оцінених параметрів моделі в майбутньому для своєчасного отримання знань, необхідних для вивчення та аналізу поведінки даних процесів.

Прийнято вважати, що майже всі процеси в природі (економічні, фінансові, біологічні, екологічні, соціальні і т.д.) є нелінійними і нестационарними. Тому процес поширення коронавірусу теж припустимо таким. Для опису таких процесів використовують відповідні математичні моделі, найбільш поширені з них авторегресія порядку p (AR(p)), авторегресія з ковзним середнім (АРКС(p, q)), моделі, які враховують наявність тренду і будуть використані в даній роботі.

Для оцінки параметрів відібраних моделей було обрано теж найбільш відомі методи: метод найменших квадратів (МНК), рекурсивний метод найменших квадратів (РМНК).

У першому розділі розглянуто актуальність розглянутої проблеми, проаналізовано історію виникнення, поширення найбільш епідемій світу та методи боротьби з ними. Також висвітлюються особливості розвитку процесів розповсюдження пандемій та актуальність задач їх аналізу.

У другому розділі розглянуто загальну методику побудови моделей процесів довільної природи на основі статистичних даних, описано вибрані для роботи моделі та методи оцінки параметрів математичних моделей, критерії адекватності для побудованих моделей та критерії якості для короткочасного прогнозу процесів, які будуть використані.

У третьому розділі побудовано та оцінено за обраними критеріями якості найкращу модель для прогнозування часового ряду даних (тих, хто захворів та тих, хто вже одужав) таких країн: України, Китаю, Іспанії, Бразилії, США. Також описано та проаналізовано декілька варіантів симуляцій процесу поширення вірусу, створених в програмі AnyLogic. На основі отриманих результатів вдалося перевірити декілька висунутих гіпотез, щодо характеру поширення епідемії за різних умов.

У четвертому розділі проведено функціонально-вартісний аналіз процесу створення програмного продукту, який втілює в собі можливість побудови вибраних математичних моделей та оцінки їх прогнозу за визначених критеріїв.

РОЗДІЛ 1: ОСОБЛИВОСТІ РОЗВИТКУ ПРОЦЕСІВ РОЗПОВСЮДЖЕННЯ ПАНДЕМІЙ ТА ЇХ МОДЕЛЮВАННЯ

1.1 Актуальність проблеми, характеристика розвитку процесів поширення пандемій в Україні та світі

COVID-19.

Спалах захворювання розпочався у грудні 2019 року у китайській провінції Ухань. 11 березня 2020 року ВООЗ оголосив його пандемією.

Пандемія COVID-19 на даний момент представляє собою серйозну глобальну загрозу здоров'ю населення. За станом на 7 червня 2020 р у всьому світі було зареєстровано 7 010 417 випадків захворювання (див. рис. 1.1):

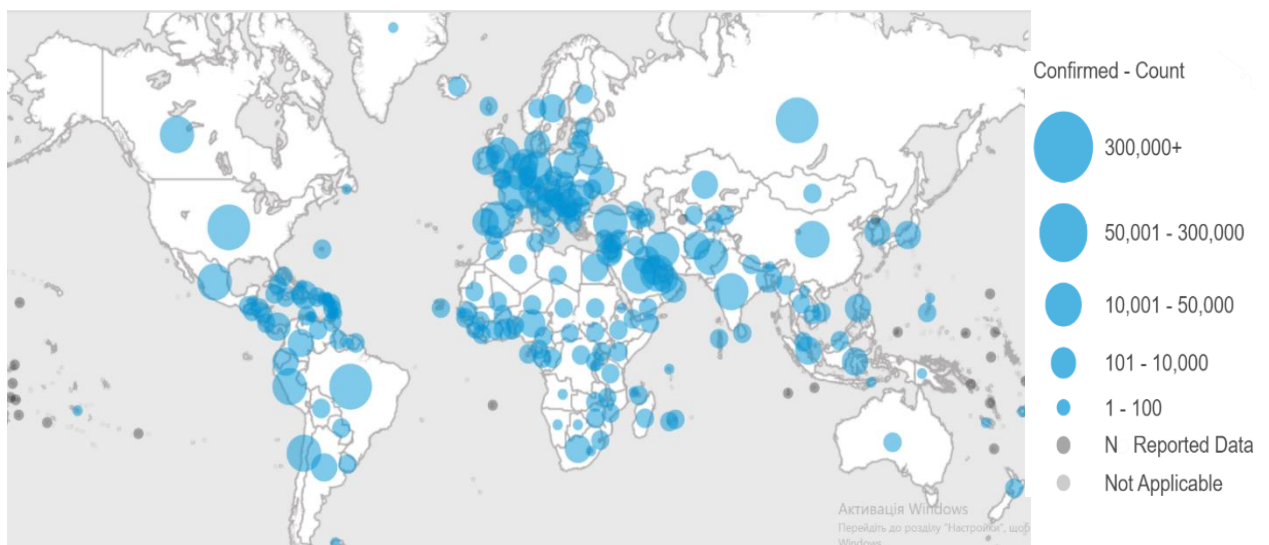


Рисунок 1.1 - Карта світу з даними про кількість інфікованих

і 404 355 смертей (див. рис. 1.2):

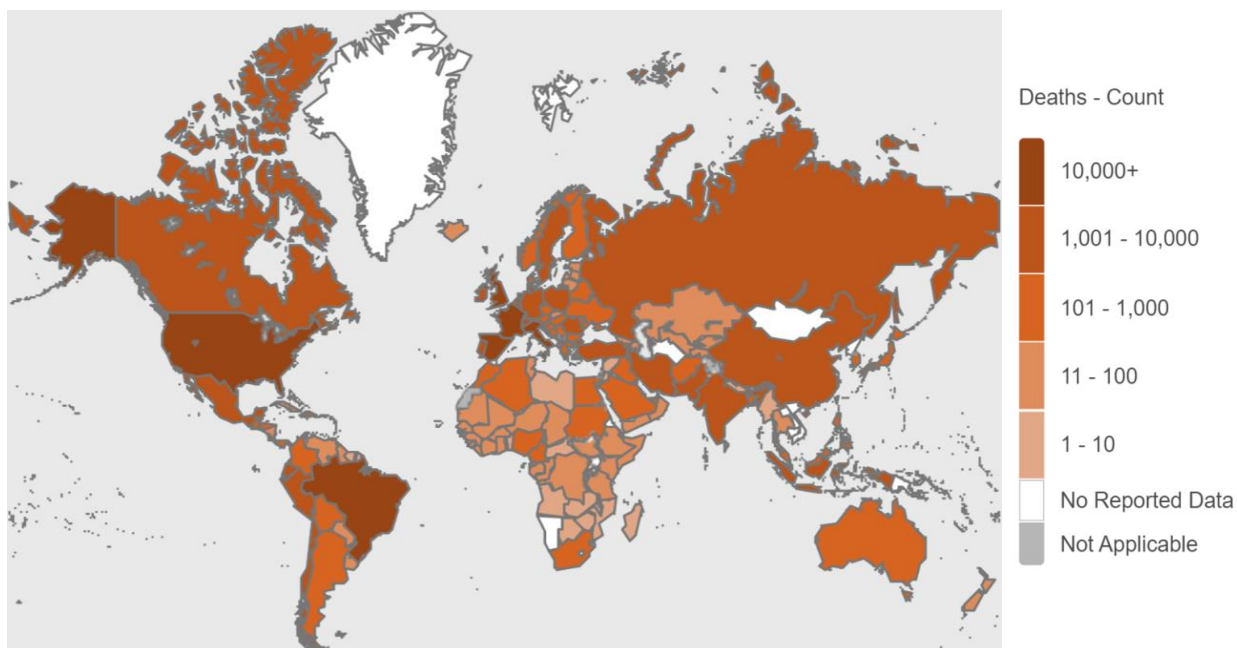


Рисунок 1.2 - Карта світу з даними про кількість тих, хто помер

Глобальне поширення відбувається швидкими темпами, і на сьогодні в кожній країні світу зареєстровано, принаймні, один випадок захворювання.

Останнього разу світ стикався з глобальною пандемією захворювання масштабів нинішньої пандемії COVID-19, що не має доступу до вакцин, - це пандемія грипу H1N1, в народі назвали Іспанка, у 1918-19 рр. В ході цієї пандемії деякі громади, зокрема в Сполучених Штатах Америки (США), відреагували на неї різноманітними нефармацевтичними заходами (НФМ) - заходами, спрямованими на скорочення передачі вірусу шляхом зниження рівня контактів серед населення в цілому. Приклади заходів, які були прийняті в той час, включали закриття шкіл, церков, барів та інших соціальних об'єктів. Міста, в яких ці заходи здійснювалися на ранніх етапах епідемії, були успішними в плані скорочення числа випадків. В принципі, доки такі нефармацевтичні міри залишалися в силі, доти в цілому була тенденція нижчої смертності. Проте, як тільки були зняті обмеження, передача інфекції відновилася.

Хоча наше розуміння інфекційних захворювань та їх профілактики в даний час сильно відрізняється від того, що було в 1918 році, більшість країн

світу стикаються сьогодні з тією ж проблемою - з вірусом COVID-19, який у 1918 році був би на одному рівні по летальності з іспанкою (грипом H1N1).

Світова пандемія призвела до серйозних глобальних соціально-економічних наслідків, відтермінування чи скасування спортивних, релігійних та культурних подій та поширення побоювань утворення дефіциту на різні товари та продукти, що в певних регіонах спричинили масову істерію і різке збільшення попиту на певні товари. Дезінформація та конспірологічні теорії про вірус поширювались інтернетом. У низці країн спалахнули протести та спостерігалися випадки ксенофобії і расизму проти китайців та інших жителів Східної та Південно-Східної Азії.

Пандемія COVID-19 уже спричинила серйозні зміни в системі охорони здоров'я та економічному добробуті багатьох країн. Вірус сильно вдарив по країнах із високим рівнем доходу, тож наслідки для слабких держав будуть іще глибшими. Надто коли йдеться про регіони, що постраждали від різноманітних конфліктів. Там наслідки довготривалої пандемії зможуть викликати найбільше проблем. Так, у значній частині Близького Сходу, де досі тривають бойові дії, вже зафіксовано випадки захворювання. На думку аналітиків, спалах коронавірусної інфекції посилить основні проблеми регіону й погіршить катастрофічну гуманітарну кризу, викликану проксі-війнами в Сирії, Ємені та Лівії. Упродовж років цивільне населення цих регіонів зазнавало постійних лих. Ще один удар у вигляді військової ескалації чи поширення вірусу стане руйнівним.

За офіційними даними, опублікованими сирійським міністерством охорони здоров'я, в Сирії зафіксовано лише 44 випадки COVID-19, три випадки смерті та 27 одужань. Однак експерти зазначають, що насправді інфекція поширеніша, ніж це визнає уряд Башара Асада. Згідно з інформацією сирійської організації, яка стежить за порушеннями прав людини, Дамаск ухвалив закриття установ і нічну комендантську годину для попередження поширення вірусу з моменту появи першого випадку. Також було заборонено переміщення громадян (винятки становили лише військові служби й

машини). Справді, бойові дії затихли, особливо після того, як Туреччина уклала угоду про припинення вогню з Росією, головним військовим союзником Асада. Однак сподіватися на довготривале перемир'я в Сирії не варто. Найімовірніше всі сторони сирійського конфлікту будуть експлуатувати тему пандемії та боротьби з нею у своїх інтересах.

Зростанню шансів на поширення вірусу може сприяти також постійна присутність турецьких і російських військових. Водночас найбільшою проблемою залишається крихка система охорони здоров'я в Сирії.

Чи були передумови для виникнення COVID-19 і які? Чому він,
найімовірніше, не останній?

Пандемію COVID-19 не можна назвати «чорним лебедем» наших днів. (Згідно з визначенням Насіма Талеба, автора концепції «чорних лебедів» – рідкісних та непрогнозованих подій з глобальними наслідками, «чорний лебідь» має бути абсолютно непередбачуваним, зокрема для експертних кіл.)

Чому ж експертні кола були такі впевнені щодо можливості пандемії
вірусного захворювання, спричиненої РНК-вірусом?

Активне масове сталие просування цивілізації у тропічні ліси з виникненням стихійних ринків диких тварин, з худобою, що супроводжує людей, та фрагментація, знищення і заміна природних біотопів пасовищами та плантаціями – все це наближує природні резервуари вірусів до людини і надає РНК-вірусам чудовий майданчик для еволюційних експериментів та освоєння «нових горизонтів» (РНК-віруси – адаптивні, здатні до дуже швидких перебудов генома, такі, що не бояться природного довкілля, високовірулентні та мінливі).

Дика природа так само глобалізується та гомогенізується. Тропічні види адаптуються у північних широтах. Лісові тварини адаптуються до життя у

сільськогосподарських ландшафтах та містах. Перебудовуються трансконтинентальні міграційні потоки птахів. Зникають або критично фрагментуються певні природні ландшафти, вимирають цілі кластери живих організмів. На їхнє місце приходять нові, адаптивні, ті, що блискавично еволюціонують, здатні пристосовуватися до людини та шалених змін навколо, які вона спричиняє. І їхні віруси так само блискавично еволюціонують.

Якщо у розвинених країнах основна частка нових інфекцій – це штами бактерій, що стали стійкими до антибіотиків, то у південних країнах, що розвиваються, велику загрозу становлять РНК-віруси, які через високі показники заміни нуклеотидних послідовностей у геномі та слабку здатність протистояти мутаціям – дуже змінні, а отже – адаптивні. А це означає, що вони потенційно здатні змінювати хазяїв та переносників.

З 1940 по 2008 роки зафіксовано виникнення 335 інфекційних хвороб, і з 1980-х років спостерігається різкий тренд на зростання їхньої кількості.

(60% – це зоонози, хвороби, що передаються від тварин до людей, і більшість з них – близько 70%, – приходять до людей з дикої природи.) Зараз 22% новітніх інфекційних захворювань асоційовано зі зміною ландшафтів та зміною землекористування.

Міжнародна торгівля та трансконтинентальні сполучення підкидають хмиз у це палаюче багаття інфекційного поширення. Певні дослідження показують, що головну загрозу для розповсюдження нових вірусів для людей становлять не стихійні ринки диких тварин у країнах третього світу, а саме глобальна торгівля дикими тваринами. Спалах пташиного грипу у Південно–Східній Азії у 2004 році почався, коли 5000 живих курчат перевезли зі Лхаси у Тибеті до Ланчжоу у Китаї, на відстань 1600 км. Лихоманку Рифтової долини завезли до Ємену з Африки разом з великою кількістю домашньої худоби, якою активно торгують між країнами у цьому регіоні. Міжконтинентальне сполучення займає кілька годин, рознесені між регіонами та країнами ланцюжки виробництв і скажені потоки імпорту та експорту, зокрема

сільськогосподарської продукції, заодно подорожують сотні видів випадкових пасажирів з вірусами «на борту».

Глобальна стрімка урбанізація сприяє перенесенню разом з населенням та домашніми тваринами зоонозів з сільської місцевості до перенаселених міст. Особливо у країнах «третього світу», це сприяє виникненню та поширенню небезпечних захворювань. У Кіншасі, столиці Конго, останніми десятиліттями рівень захворювання на раніше ендемічну сонну хворобу зріс у десятки разів через потужну міграцію сільського населення до міста. У Кабул (Афганістан) разом з новими мешканцями з сільськогосподарських районів у 1987 році повернувся шкірний лейшманіоз, до якого аборигенне населення мало імунітет.

Всю свою історію людство боролось проти смертоносного і невидимого ворога. Випадки масового зараження описані навіть в Біблії. З чим чи з ким саме ми воювали? Бактерії, віруси... Але звідки вони підкрадалися завжди непомітно, завжди неочікувано? Я провела дослідження усіх найбільш смертоносних пандемій світу, які пам'ятає людство і проаналізувала момент виникнення збудника захворювання, що саме посприяло масштабному ураженню населення і як люди в свій час з цим боролися.

1.2 Історія виникнення та поширення найбільших пандемій у світі до 20 ст.

1). Юстиніанова Чума (в честь імператора Візантії Юстиніана 1)

541-750 рр.

Перша в історії пандемія, яка захопила весь цивілізований світ: Європа, Північна Африка, Центральна і Південна Азія, Аравія.

Природними резервуарами для хвороби стали гризуни – щури, миші, ховрахи, піщанки з Центральної Азії, на яких неодмінно паразитували блохи – носії бактерії чумна паличка.

Торговими шляхами разом з караванами з шовком та іншим крапом подорожували і тварини, переважно щури. Спочатку в Ефіопію чи Єгипет, потім середземноморськими торговими каналами у Константинополь (загинуло 66% населення), який був важливим великим торгівельним вузлом-центром з найсприятливішими умовами зараження і поширення захворювання серед людей.

Кількість загиблих: близько 66 млн. осіб на Сході, до 25 млн. осіб в Європі.

2). Чорна смерть або бубонна чума

1347-1353 рр.

Як і першого разу збудником хвороби була бактерія чумна паличка, яку переносили блохи, які в свою чергу паразитували на гризунах. Тільки тепер поширенню посприяли не тільки «Шовковий шлях» через Центральну Азію, а й міжусобні війни монголо-татарських феодалів за право бути ханом Золотої Орди, яка вже почала розпадатися, на тій території приблизно в 1330-х роках. До Західної Європи бубонна чума потрапила з Криму, припливши на генуезьких кораблях наприкінці 1347 року.

Чума поширювалась в основному в містах, які були переповнені людьми, щурами і потопали у власних нечистотах. Села, в порівнянні з містами, були ізольовані, тому більш захищені від можливості заразитися. Згадується, що жителі деяких сіл спеціально закривали поселення на в'їзд.

Методи боротьби: пастеризація для населення(дезінфекція продуктів і рук оцтом, вином, пивом); спеціальні костюми зі шкіри з масками з травами для лікарів, знахарів, священників, монахів; введення карантину.

Перший карантин (перекл. 40 днів) тривалістю якраз 40 днів ввели у Венеції (на той момент центр середземноморської торгівлі) для всіх кораблів

та екіпажів, які прибували з уражених територій (карантинну зону організували поблизу на о. Лазоретто). Також у місті заборонили торгівлю вином і азартні ігри, закрили борделі, списували боржникам частину суми, звільняли людей з в'язниць, наймали людей прибирати мертвих з вулиць та будинків.[6]

Кількість загиблих: за 10 років Європа до 20-40% населення, Китай та Індія – вимерло декілька сотень міст і сіл, Африка – неможливо підрахувати.

3). Портова чума

1894-1921 рр.

1855р., Китай. Гірські долини Юньнаня хоч і були природнім осередком чуми, де протягом 4 десятиліть зустрічались випадки захворювання, були доволі ізольовані від зовнішнього світу, незважаючи на підвищений інтерес в другій половині 18 ст. до регіону через відкриті корисні копалини, зокрема мідь (кількість населення тоді зросло понад 7 млн.). Відповідно зросла кількість контактів з дикими жовтогрудими пацюками, основними збудниками чуми в тому регіоні.[7]

Паралельно на порозі 20 ст. коли довірені імператора купці активно ведуть торгівлю з Європою, в Китай попадає опіум. Введення суворих обмежень на вільну торгівлю китайським імператорським урядом з надією побороти поширення наркоманії призводить до двох опіумних війн з Британією (дві поразки). Чим користуються місцеві революціонери і організовують державний переворот.[2]

Як наслідок, на початку 20 ст. революція, масові повстання селян і солдат в країні, безпосередня близькість до природних резервуарів призводять до стрімкого поширення чуми.

Кількість загиблих: загалом третя хвиля чуми забрало життя близько 20 млн. людей Китаю та Індії.

Загальна кількість життів, що забрала чума за 100 років більше 150 мільйонів при населенні 400 – 450 млн. всього. Розвиток міжнародної торгівлі,

загальна урбанізація, при цьому абсолютна неготовність міст прийняти таку кількість людей, повне ігнорування будь-якої гігієни, відсутність елементарних комунальних систем для життєздатності і забезпечення першим необхідним міст. Ми вчимося на власних помилках, але платимо за це занадто високу ціну.

4). Холера

Смертельна кишкова інфекція, яку спричиняє бактерія холерний вібріон, відома людству з часів Гіппократа і Цельса. Вважається, що стародавні прирічкові культури Індії послужили розвитку патогенних властивостей у водних сапрофітів-вібріонів, які в свою чергу перетворилися вже у холерні мікроби. Тобто, в наслідок природного відбору бактерії в річковій воді змогли пристосуватися до паразитування в кишечнику людини. Ті з них, що викликали діарею, забезпечували швидке поширення захворювання, тим самим продовження собі життя.

Отож, спалахи захворювання виникали, але доволі локально і тільки на території Індії. Цьому сприяли висока щільність заселення берегів основних водних артерій країни, їх постійне фекальне забруднення, що унеможлиблювало безпечність питної води, повна антисанітарія. Військові та паломницькі походи, каравани могли посприяти масштабному поширенню хвороби по світу, але бактерія не могла адаптуватися за межами Індії через несприятливі кліматичні умови до 1816 року. [3]

Початок 19 ст. – це промислова революція, стрімке зростання міжнародної торгівлі, чисельність колоніальних воєн, значна міграція людей, ослаблених постійними епідеміями (тільки закінчилась портова чума, ведеться боротьба з натуральною віспою(масова вакцинація), грип).

- 1816-1824 рр. – **Перша холерна пандемія.**

Території ураження: Південно-східна Азія, Близький схід, Східна Африка, середземноморське узбережжя Європи.

Кількість загиблих числиться сотнями тисяч, можливо мільйонами. В тому числі постраждало багато британських солдатів, що й привернуло увагу Європи. [4]

- 1826-1837 рр. – **Друга холерна пандемія.**

Територія ураження: вся Східна Азія (чого не відбулося першого разу), через Західну Азію до Європи та Америки, також на схід до Китаю та Японії. У 1830 р. вперше занесено в Україну під час російсько-османської війни. У Російській Імперії виникають «холерні» бунти через спроби влади ввести суворі карантинно-поліцейські заходи.

Кількість загиблих невідомо.

- 1846-1860 рр. – **Третя холерна пандемія.**

Територія ураження: поміж раніше залучених територій вперше поширилась на всю Північну Африку і Америку.

Розповсюдженню хвороби сприяли паломництво шіїтів, мусульман до священних місць, пілігрими, які з ними контактували та критично мала кількість громадських вбиралень, рої мух в місцях скупчення людей.

У 1847 р. було охоплено всю Україну, Польщу, Російську імперію. Дісталась Великої Британії, де вперше довели, що зараження відбувається через воду.

За кількістю загиблих цю пандемію вважають найбільшою серед усіх в 19 ст.

- 1863-1875 рр. – **Четверта холерна пандемія.**

Відкриття Суецького каналу посприяло легкому та швидкому перевезенню хвороби з Індії до Європи (швидкість проникнення скоротилась з двох, трьох років до декількох тижнів). Далі естафету підхоплювали залізничні сполучення і четверта хвиля запам'яталася стрімким поширенням хвороби.

- 1881-1896 рр. – **П'ята холерна пандемія.**

Вперше початок пандемії був не в Індії, а в Єгипті, під час якої було доведено бактеріальну природу хвороби і виділено збудника *V. cholerae*.

На цей раз він потрапив і в Південну Америку.

При тяжкій формі захворювання після смерті спостерігаються швидкі спазми усіх груп м'язів, людина ще може рухатись і навіть сісти. Думки про те, що ховають живих, серед народу підняли масові заворушення на уражених територіях.

- 1899-1923 рр. – **Шоста холерна пандемія.**

Територія ураження: всі континенти, окрім Атарктиди. Поширенню посприяла Перша світова війна і все, що за нею послідувало.

В 1905 році було виявлено новий збудник холери *V. cholerae* біовар El Tor зі смертністю близько 70%, на відмінну від класичного збудника – 44-50%, при цьому до 1961 року цей біотип вважали абсолютно нешкідливим.

- 1961-1970 рр. – **Сьома холерна пандемія.**

На цей раз винуватцем був *V. cholerae* біовар El Tor. Свого максимуму досягла в 1970 р. і продовжується дотепер. Було знайдено декілька нових штамів, серед яких знову ж зустрічаються з вищою летальністю, як в класичного збудника.

За період незалежності України неодноразово виявляли холерні вібріони в басейнах Чорного, Азовського морів, Дніпра, Дністра, Південного Бугу. В 2011 р. у Маріуполі ввели заборону на купання, виловлювання риби та продаж її на ринках.[5]

За свою столітню історію холера забрала більше 10 млн. життів за доволі оптимістичними підрахунками. Навіть у сучасному світі час від часу реєструють нові спалахи холери.

Проте в пошуках ефективних методів боротьби з напастю цивілізований світ нарешті визнав, як необхідність створення та подальшого вдосконалення централізованої каналізаційної системи і водопостачання в містах, фільтрувальних станцій, інститутів гігієни, управління охорони здоров'я.[6]

Так, ми навчилися з нею боротися: лікувати, прогнозувати, враховуючи кількість опадів, нові потенціальні осередки за декілька тижнів до їхнього формування, понизили смертність до 0.21%, але цього все одно недостатньо. До поки у світі житимуть люди за межею бідності, в повній антисанітарії, без доступу до безпечної питної води, доти ми не зможемо віднести холеру до списку забутих людством хвороб.

5). Натуральна віспа

6 – 20 ст.

Одне з найперших описаних людством захворювань ще за Стародавніх Греції і Риму. Збудником віспи є *Strongiloplasma variola majoris* — досить великий вірус розмірами 210–260 нм. Його сліди були виявлені на зразках тканин мумій Стародавнього Єгипту. [8] Вірус вирізняється унікальною стійкістю до будь-яких кліматичних умов, температур, речовин, які мають дезінфікуючі властивості, антибіотиків.

Загальна смертність: за оптимістичних підрахунків щонайменше 1 млрд. людей. Тільки за 20 ст. припускається, що це захворювання убило до 300 млн. людей (це більше ніж в усіх війнах людства).

Першу успішну безпечну вакцину від віспи розробив англійський лікар Е. Дженнер у 1796 році. Помітивши таку закономірність: люди, які безпосередньо контактують з тваринами (доярки, пастухи, кавалеристи) тільки зрідка можуть заразитися. Виявилось, що перехворівши коров'ячою віспою, у людини виробляється стійкий імунітет до натуральної.

Єдине повністю ліквідоване у природі шляхом вакцинації захворювання у 1980 році за офіційним повідомленням ВООЗ. Незважаючи на оголошений факт деякі країни продовжують виробляти протівіспові вакцини, ВООЗ створила її запаси, а уряд США відновив щеплення для військовослужбовців з метою мінімізації ризику використання вірусу, як біологічної зброї масового ураження та забезпечити гарантовану готовність до будь-якої надзвичайної ситуації у сфері охорони здоров'я.

Це мало не єдине захворювання протягом усієї історії людства і до 20 ст., яке викликало масштабні пандемії по всьому світу, забрало життя рекордної кількості людей (а ті одиниці, що вижили залишились в більшості випадків сліпими, каліками) і збудником якого являється вірус, а не бактерія.

Це доволі цікава закономірність... І не згадані вище захворювання, що спричиняли епідемії, пандемії, до 20 ст. – бактеріальні:

- **Тиф** (збудником є бактерія тифу;
вперше описаний випадок епідемії в Афінах в 40 р. до н.е.;
тільки під час Першої світової війни в Україні, Польщі та Росії померло 3.5 млн. осіб; на сьогодні смертність становить 5-8%)
- **Малярія** (збудником є малярійний паразит, вік якого оцінюють біля 600 млн. років; тільки за 2000 р. захворіло 350-500 млн. людей, з них до 3 млн. летальних; першу успішну вакцину придумали тільки в 2017р.)
- **Туберкульоз або Біла чума** (в Україні наз. сухоти) (збудником є паличка Коха; виявлено навіть у єгипетських мумій; 1600-ті – 1800-ті роки – найбільша епідемія туберкульозу, помирав кожний 7-мий інфікований; в наш час від цієї бактерії щорічно вмирає 1.8 млн. осіб) [9]
- **Проказа** (збудник бацила *Mycobacterium leprae*; перші згадки в Китаї, Єгипті, Індії в 600 р. до н.е.; в боротьбі з захворюванням вперше використані були карантинні ізолятори – лепрозорії; з 1985 року вилікували 15 млн. людей).

1.3 Історія виникнення та поширення найбільших пандемій у світі 20-21 ст.

1). Іспанка (іспанський грип)

Січень 1918 – Грудень 1919 рр., штам H1N1

Гостре інфекційне захворювання викликане вірусом грипу А. Вдалося встановити, що природними носіями для більшості різновидів цього серотипу грипу (за вірусологічною класифікацією вірус грипу є РНК-вірусом і включає три серотипи А, В, С) є дикі водоплавні пташки.

Кінець Першої світової війни. В необтяженій цензурою воєнних часів іспанській газеті з'являються заголовки про спалах нової невідомої хвороби. Як пізніше з'ясується джерелом вірусу був польовий табір воєнних у Франції, через який проходило 100 000 солдатів за день, поруч знаходилась свиноферма і регулярно привозили домашню птицю.

За 1 рік захворіло близько 500 млн. осіб (29.5% населення планети), починаючи з Європи та Північної Америки і до найвіддаленіших куточків планети. Пандемія іспанки затихла до весни 1919 року, вірус зник так само несподівано, як і з'явився. Існує теорія, що штам грипу дуже швидко мутував до менш летального.

В зоні ризику були здорові люди віком 18 - 40 років.

Кількість загиблих: за приблизними оцінками 50-100 млн. осіб (5% населення планети).

Після найсмертельнішої пандемії в історії спостерігаються хвилями спалахи епідемій грипу:

- 1957-1958 рр. - Азіатський грип (штам H2N2)
Кількість загиблих: приблизно 70 000 людей.
- 1968-1969 рр. – Гонконгський грип (штам H3N2)
Кількість загиблих: приблизно 1 млн. людей
- 2003-2008 рр. – Пташиний грип (штам H5N1)
Кількість загиблих: приблизно 360 людей
- 2009-2010 рр. – Свинячий грип (штам A/H1N1)

Кількість загиблих: по оцінкам експертів від 151 700 – 575 400 людей, лабораторно підтверджених тільки 18 632. 80% смертей серед осіб молодше 65 років.

Цікаво, що передувала всім цим подіям ситуація циклічного поширення вірусу-попередника у європейських арміях за декілька місяців, можливо років до. Цей факт, а також те, що переносниками вірусу були птахи, які в результаті мутації змогли заразити свиней зі свиноферми близько фронту, стверджувались Оксфордом та його командою і паралельно згадувались в доповіді, опублікованій «Journal of the Chinese Medical Association» в 2016 р. В 2017 році історик Сантьяго Мата оголосив, що на кінець 1917 року була зафіксована перша хвиля епідемії принаймні в 14 військових таборах США.

2). ВІЛ/СНІД

1981 – сьогодення

Вперше він з'явився в Центральній Африці. Раніше ним хворіли тільки мавпи ще в 1920-х роках.

5 липня в 1981 в США було зафіксовано 5 випадків інфікування. До кінця 1982 року виявили 771 випадок, з яких 618 померли ж в тому році. Тільки тоді цивілізований світ вперше серйозно заговорив про це захворювання і необхідність проведення вірусологічних досліджень.

На даний момент ВІЛ-інфекція зареєстрована у всіх країнах світу.

Вакцини від цього вірусу на даний момент не існує, бо впродовж розвитку ВІЛ-інфекції в однієї людини вірус постійно мутує і утворює нові штами стійкі до ліків, які приймає пацієнт. Тому захворювання рахують невиліковним, хоча антиретровірусною терапією можна підтримувати вірус в пригніченому стані.

Кількість інфікованих: приблизно 40 млн. ВІЛ-інфікованих.

Кількість загиблих: 35 млн. осіб.

3). Ебола

1976 – сьогодення

Вперше зафіксували випадки геморагічної лихоманки Еболи в екваторіальній провінції Судану в 1976. Тоді захворіло 602 людини, з них

померло 331. Природними носіями вірусу Еболи є крилани (лат. Pteropodidae, різновид кажанів). Вразливими до інфікування є деякі примати, парнокопитні, відомі випадки зі свинями і дукерами (підвид антилоп), люди. Припускають, що мешканці Африки інфікувалися під час полювання на мавп, невеликих тропічних копитних, які підхоплюють вірус від кажанів. З моменту виявлення майже кожного року реєструються нові випадки, але через високу летальність і короткий інкубаційний період не встигає поширитися. Проте завдяки особливості консервувати себе в клітинах, вірус знаходиться в тканинах організму певний період після видужання, що можливо повпливало на спалахи епідемій, хоч і досить локальних:

- 1995 р. ДРК, 315 хворих, з них померло 254 (летальність 81%).
- 2000-2001 рр. Уганда, 425 хворих, з них померло 224 (летальність 53%).
- 2007 р. ДРК, 264 хворих, з них 187 померло (летальність 71%).
- 2014-2015 рр. захворіло 27748 осіб, з яких померло 11279 (летальність 41%). Епідемія поширилась в Західній та Центральній Африці, США, Великобританії, Іспанії, Італії, Малі.
- 2018-2020 рр. ДРК, захворіло 3366 осіб, з них померло 2227 (летальність 66%).

Вірус Еболи мутує і на сьогодні відомо мінімум 5 його штамів. Ніхто не може бути впевнений у тому, що наступна мутація не дасть йому можливість спростити шлях передачі до людей і викликати пандемію в світі, перетворивши вірус Еболи з екзотичної хвороби Африки на проблему для людства, як це сталося з ВІЛ/СНІД, малярією і т.д.

Наприклад, нещодавно у Філіппінах знайшли новий штам вірусу, який уражає і людей, і домашніх свиней.

4). SARS/MERS

SARS (2002 р.) - вірус «атипової пневмонії» з'явився у китайській провінції Гуаньдунь, Південний Китай у 2002 році. Збудником є коронавірус

SARS-CoV. Інфекція виникла на «живому ринку» від китайської цівети, яка спрацювала, як проміжний господар або господар-підсилювач. Природнім переносником коронавірусу був кажан з роду підковиків. Під час епідемії було виявлено близько 8000 підтверджених випадків, з них приблизно 800 летальних.

MERS (2015 р.) – близькосхідний респіраторний синдром, виникає при ураженні коронавірусом типу MERS-CoV. Перший випадок зафіксований на початку осені 2012 року у туриста, який приїхав з Саудівської Аравії в Британію. Вторинним резервуаром виявились верблюди, а от первинного до цих пір не знайшли. Встановлення способу передачі людині цього вірусу зіграє вирішальну роль в боротьбі проти нього.

5). COVID-19

2019-2020 рр.

Уперше коронавіруси були описані в 1965 році. Причинами виникнення нових і нових штамів таких, як SARS-CoV або MERS-CoV, є постійні численні мутації (адаптації вірусу до навколишніх умов). Остання коронавірусна інфекція яляється зооантропонозною. Джерелом захворювання можуть бути і дикі, і домашні тварини (птахи, мавпи, змії та ін.). У випадку з COVID-19 можливими природніми резервуарами виступили кажани, які переносили збудник SARS-CoV-2.

Значення інфікованих підтверджених випадків доходить до 5 млн. осіб, з них померло уже більше 310 000 людей.

6). CCHFV (Вірус геморагічної лихоманки Крим-Конго)

Вперше наш потенційний український герой виявлений в Криму у 1944 році. Природнім носієм і переносником являється іксодовий кліщ роду *Hyalomma* через укуси або з кров'ю домашніх тварин, раніше ним інфікованих. Вірусологи побоюються двох можливих варіантів розвитку подій:

Для даного кліща оптимальними кліматичними умовами є дуже тепле літо і волога зима. Глобальні зміни клімату можуть відсунути межу далі на північ, а перелітні птахи зіграють роль кур'єрської доставки.

Вірус мутуватиме, поки не перекинеться на нового господаря, будь-то просто більш холодолюбний іксодовий кліщ-родич чи хтось з переносників кліщових енцефалітів.

Зміна сільського господарства супроводжує зміни клімату, і багато епідемій починають розширювати ареал і просуватися на північ. Вірус Рифтової долини (RVF), лихоманки Денге та Чикунгунья – особливо яскравий приклад швидкого поширення вірусних тропічних небезпечних хвороб у нові місця, зокрема до благополучної Європи. Всі ці хвороби асоційовані з комарами, кажанами і т.д., а отже, всі безпосередньо пов'язані зі зміною зволоження, кліматичних умов та створенням нових придатних для цих тварин біотопів в безпосередній близькості з людиною – сільськогосподарських угідь, підвалів будинків тощо.

1.4 Принципи моделювання епідемії

Багато наукових команд використовують унікальні моделі, що роками, а іноді й десятиліттями розроблялись та вдосконалювались у відповідному дослідницькому центрі. Однак математичні принципи побудови усіх цих моделей є дуже схожими. Всі вони будуються для того, аби краще зрозуміти, як саме та наскільки швидко люди рухаються між трьома можливим станами: вразливі до вірусу (S), інфіковані (I), ті, що одужали (R) або померли. Припускається, що група R має імунітет до вірусу і більше не може передавати інфекцію. Також до деяких вірусів у певних людей може бути природний імунітет. При моделюванні таких людей також відносять до групи R.

Окремі епідеміологічні моделі розроблялися протягом майже століття і добре апробуються на даних минулих епідемій. Ці моделі засновані на моделюванні фактичного процесу зараження (трохи схоже на прогнози погоди, що моделюють атмосферні процеси). Найпростіші моделі засновані на класифікації живих індивідуумів в популяції як сприйнятливих, інфекційних або рекуперованих - тому вони називаються SIR-моделями. Найпростіші моделі SIR роблять базові припущення, як-от наприклад, що всі мають однаковий шанс підхопити вірус, а також, що інфіковані люди в однаковій мірі передають вірус протягом всього періоду захворювання. Більш складні моделі, на яких власне і ґрунтують свої рішення політики під час пандемії, поділяють людей на малі групи – за віком, статтю, станом здоров'я, кількістю контактів тощо, аби краще зрозуміти, хто зустрічається з ким, коли та в яких місцях.

Такі більш складні варіації дозволяють класифікувати ще кілька категорій індивідуумів. Ці моделі допускають як недіагностовані, так і діагностовані випадки, тому вони можуть враховувати обмежене тестування. Наявні дані використовуються для оцінки різних параметрів в моделях, а потім проводиться моделювання, щоб побачити, що буде відбуватися за різними сценаріями. Деякі з параметрів можуть бути тільки грубо оцінені, але ми можемо перевірити, що станеться з прогнозами, варіюючи параметри в межах можливого діапазону.

Використовуючи детальну інформацію про щільність населення, частку людей похилого віку, транспортне сполучення, розмір соціальних груп, стан системи охорони здоров'я та інші фактори, математики будують віртуальну копію міста, регіону або ж цілої країни за допомогою диференційних рівнянь, що відображають взаємодію різних груп населення у просторі та часі. Потім у цей віртуальний світ додають вірус і «спостерігають» за розвитком подій.

Однак у процесі моделювання, вченим доводиться робити досить багато припущень щодо тих чи інших показників розповсюдження вірусу. Наприклад, необхідно оцінити рівень смертності та базове репродукційне число (R_0) – кількість людей, яким в середньому передає вірус інфікована

особа. Наприклад, у своєму звіті від 16 березня команда Імперського коледжу Лондона припустила, що для конкретної демографії Великої Британії рівень смертності від COVID-19 становить 0.9%. Також, базуючись на доступній інформації з різних країн світу, науковці припустили, що одна людина в середньому передає вірус двом-трьом іншим людям (R_0 знаходиться в межах від 2 до 2.6), що вірус SARS-CoV-2 потребує близько 5 днів для інкубації в інфікованій людині, а також зробили низку інших припущень щодо ймовірності передачі хвороби на тій чи іншій стадії захворювання.

Методи прогнозування інфекційних захворювань активно розвиваються з 20 ст. Останні декілька років завдяки створенню інформаційних систем наглядку, появи великої кількості статистичних даних, доступних для аналізу, число робіт, пов'язаних з цією темою, стрімко зростає. Наприклад, доктор Елісон Хілл (Гарвард) створила відмінний додаток Shiny, який дозволяє досліджувати вплив різних параметрів в описаних моделях.

А взагалі першопрохідцями в моделюванні процесу поширення інфекційних захворювань були: Вільям Гамер та Рональд Росс, який на початку XX століття застосував закон дії мас, щоб пояснити поведінку епідемії. Ловелл Рід і Вейд Гемптон Фрост розробили епідемічну модель Reed-Frost, щоб описати зв'язок між сприйнятливими до захворювання, інфікованими і тими що набули імунітету особинами в популяції.

Однією з базових моделей, яка була успішно дослідженою, стала модель Кермака-МакКендріка, побудована в 1927—1933 роках (оригінальні статті тих часів було передруковано в 1991 року).

Ці моделі відомі як **полігамні моделі в епідеміології**, а також слугують базовими математичними моделями, які дають змогу зрозуміти складну динаміку та основні особливості цих систем.

Прогнозувати процес поширення пандемії складніше, ніж багато хто думає. У книзі Джорджа Атанасопулоса обговорюються фактори, які сприяють тому, щоб прогнози були відносно точними.

Виділено три основні чинники:

1. наскільки добре ми розуміємо чинники, що сприяють цьому;
2. наскільки доступні дані;
3. чи може прогноз вплинути на те, що ми намагаємося передбачити.

Наприклад, завтрашню погоду можна відносно точно прогнозувати, використовуючи сучасні інструменти, тому що у нас є хороші моделі фізичної атмосфери, є тонни даних, і наші прогнози погоди не можуть вплинути на те, що відбувається насправді.

З іншого боку, прогнози завтрашніх цін на акції набагато менш точні, тому що фактори 1 і 3 вище не задовольняють. По-перше, чинники, які сприяють змінам цін на акції, не дуже добре вивчені і, в крайньому разі, частково залежать від людської психології. По-друге, добре розрекламовані прогнози фондового ринку можуть безпосередньо впливати на поведінку багатьох інвесторів.

Якщо застосувати ці ідеї до пандемії COVID-19, то легко зрозуміти, чому прогнозування її наслідків затруднено. Незважаючи на те, що ми добре розуміємо, як вона працює з точки зору міжособистісних інфекцій, ми маємо в своєму розпорядженні дані, які є обмеженими і можуть вводити в оману.

Відомо, що нинішнє число підтверджених випадків захворювання значно занижено через обмежений доступ до тестування. Майже напевно існує набагато більше випадків COVID-19, які не були виявлені, в порівнянні з підтвердженими. Крім того, рівень недооцінки сильно варіюється між країнами. У такій країні, як Південна Корея, де багато тестувань проведено, кількість підтверджених випадків буде ближче до кількості реальних випадків, ніж у США, де тестування було набагато менше. Тому ми просто не можемо легко змодельовати поширення пандемії, використовуючи наявні дані.

Друга проблема полягає в тому, що прогнози COVID-19 можуть вплинути на те, що ми намагаємося передбачити, тому що уряди реагують, деякі краще, ніж інші. Проста модель, яка використовує наявні дані, буде вводити в оману, якщо вона не може включати в себе різні кроки, що робляться

для уповільнення передачі. Підгонка простих моделей під наявні дані безглузда, вводить в оману і небезпечна.

ВИСНОВКИ ДО РОЗДІЛУ

Пандемії стали двигуном медичних досліджень, посприяли розробці антибіотиків, використанню вакцинацій, які врятували згодом мільйони життів.

До 20 ст. люди не володіли достатньо ефективними методами боротьби з бактеріальними захворюваннями, при цьому умови життя були жахливими, що тільки сприяло затягненню періоду епідемій, пандемій на десятки років, століть.

З розвитком медицини, охорони здоров'я, покращенням умов життя, дотриманням правил елементарної гігієни, людство скоротило до мінімуму ризик спалаху захворювання бактеріального характеру. Ми навчилися наперед передбачати і прораховувати ймовірність такої епідемії. Як результат, за останні 100 років ефективно контролюємо і лікуємо нові зафіксовані випадки, не даючи змоги поширитися.

Історія розглянутих хвороб показує нам, що епідемії могли продовжуватися століттями, аж доки не з'являлася ефективна універсальна вакцина, або кардинально змінювалися умови у суспільстві, роблячи ризик інфекції мінімальним. «Раптові» хвороби, такі як «іспанка» особливо змушують задуматися над майбутнім розвитку COVID-19. Чи обмежиться хвороба трьохрічним періодом (захворювання попередніми штамми коронавірусу помічали у КНР та Республіці Корея з 2017 р.), чи буде продовжуватися щорічно (як відбувається з «верблюжою хворобою» (MERS-CoV) – близькосхідним штамом коронавірусу)?

Кажани, резервуари всіх цих страшних хвороб, тисячоліттями сиділи з вірусами в організмі десь у непролазних лісах Південно-Східної Азії, Конго, Уганди, в австралійському буші.

Зміна сільського господарства супроводжує зміни клімату, і багато епідемій починають розширювати ареал і просуватися на північ. Вірус Рифтової долини (RVF), лихоманки Денге та Чикунгунья – особливо яскравий приклад швидкого поширення вірусних тропічних небезпечних хвороб у нові місця, зокрема до благополучної Європи. Всі ці хвороби асоційовані з комарами, а отже, всі безпосередньо пов'язані зі зміною зволоження, кліматичних умов та створенням нових придатних для комарів біотопів – сільськогосподарських угідь, підвалів будинків тощо.

Масштабні епідемії будуть спричинені появою нових видів вірусів, до яких не пристосований людський імунітет. Очевидно, що визначальним фактором у поширенні епідемії є густина населення та соціальна активність популяції. Тому осередком всіх масштабних епідемій є великі міста. Ухань є головним транспортним вузлом Китаю з населенням більше 11 млн. осіб. Багато вірусів швидко мутують, утворюючи численні штами та підвиди. Ці фактори істотно ускладнюють роботу з їх вивчення, створення ефективних методів лікування та прогнозування поширення епідемій. Сьогодні продовжується процес урбанізації не тільки в Україні, а й у всьому світі, тому, на жаль, можна констатувати, що в майбутньому кількість та розмах епідемій не зменшуватимуться.

І якщо бути чесними самим з собою, ні наша система охорони здоров'я, ні наша політична свідомість, соціальна відповідальність, в принципі наша цивілізація можлива так ніколи і не зможе бути підготовлена до нових викликів природи.

Лише через кілька місяців, а може, й років, можна буде оцінити, наскільки точними є математичні моделі, що сьогодні намагаються спрогнозувати поширення коронавірусу, і відповідно, лежать в основі критичних рішень, які ухвалюють уряди різних країн. На теперішній момент

це найкращий доступний інструмент, але варто пам'ятати, що це лише моделювання, засноване на певних припущеннях, які, цілком ймовірно, можуть виявитися хибними.

РОЗДІЛ 2 :ВИБІР І ОПИС СТРУКТУР МАТЕМАТИЧНИХ МОДЕЛЕЙ ПРОЦЕСІВ ПОШИРЕННЯ ПАНДЕМІЇ

2.1 Підходи до математичного моделювання вибраних процесів

Епідемічний спалах, викликаний коронавірусом SARS-CoV2, викликає величезний інтерес для різного роду досліджень з погляду на високі темпи поширення та досить значну кількість смертельних випадків. Детальний науковий аналіз цього феномену паралельно проводять декілька передових університетів світу, але суспільству хочеться знати наперед про потенційно можливий період протяжності епідемії, очікувану кількість захворілих та смертей. Для довгострокових прогнозів необхідні складні математичні моделі, які потребують багато зусиль та часу для ідентифікації та обчислення невідомих параметрів.

В даному розділі будуть представлені деякі методи оцінки прогнозу поширення інфекційних захворювань, в тому числі і коронавірусу.

Метою роботи є вивчення та пошук закономірностей поширення вірусів, вивчення ефективності різних засобів боротьби з епідеміями за заданих обставин та створення системи аналізу та моделювання поширення вірусних захворювань.

Основним завданням прогнозування розвитку епідемії є не оцінка кількості людей, що захворіють, та ступенів складності захворювання, а розроблення ефективних методів стримування та зупинки епідемії. Швидкість поширення вірусних захворювань не дає змоги експериментувати зі засобами стримування, а вимагає швидкої ефективної стратегії боротьби з епідемією вже в перші дні поширення. Це можливо лише за наявності адекватної комп'ютерної моделі розвитку епідемії в умовах певної країни та можливості швидкого вивчення різноманітних засобів стримування та профілактики на поширення вірусу.

Математичні моделі, побудовані на основі диференційних рівнянь, можуть бути дуже складними, коли усе населення поділяється на досить малі групи для кращого відображення реальності. Тож існує альтернативний підхід до моделювання поширення вірусу, який дозволяє максимально детально відобразити реальний світ, уникаючи складності моделей, заснованої на рівняннях. Такий підхід базується на створенні та використанні так званих «агентів», які діють відповідно до специфічних правил поведінки окремих індивідів. Ці моделі є менш складними, але потребують дуже багато даних на рівні окремих домогосподарств: хто і як добирається до роботи, де і з ким проводить вихідні тощо.

Слід зауважити, що моделі, побудовані на диференційних рівняннях та навколо так званих агентів, дають досить схожі прогнози. Наприклад, 26 березня Імперський коледж Лондона випустив ще один звіт з глобальними прогнозами, зробленими на основі відносно простої моделі, побудованої за допомогою диференційних рівнянь. Відповідно до цього звіту, якщо в США не будуть вживати жодних заходів проти коронавірусу, то втратять 2.18 мільйона життів. Більш рання модель, побудована на базі агентів, спрогнозувала рівень смертності у США за тих самих умов на рівні 2,2 млн людей.

Враховуючи величезну швидкість поширення вірусу як в межах окремого міста, так і у всьому світі, на перший план виходить оперативність та адекватність реагування державних органів на епідеміологічну ситуацію. Епідемії останніх років яскраво показали, що прогнозування та оцінка епідеміологічної ситуації далекі від ідеалу. Експерти складають прогнози на основі обмежених та недостовірних наборів даних про аналогічні віруси, а також на підставі інформації про поширення вірусу в інших країнах. Часто такі прогнози не дають навіть близького до правди результату.

2.2 Методика побудови моделей вибраних процесів

2.2.1 Узагальнена класифікація динамічних процесів

В економіці, фінансах, екології, демографії та інших областях людської діяльності існує велика різноманітність різних процесів. Однак існують деякі загальні риси процесу, такі як лінійність / нелінійність і стаціонарність / нестаціонарність, які дозволяють розділити їх на практично зрозумілі класи і вибрати відповідні методи моделювання і прогнозування. На рис. 2.1 показана спрощена класифікація процесів, з якої можна зробити висновок про велику різноманітність структур математичної моделі, які можуть бути застосовані для формального опису динаміки процесів і рішення задачі прогнозування їх еволюції в часі.

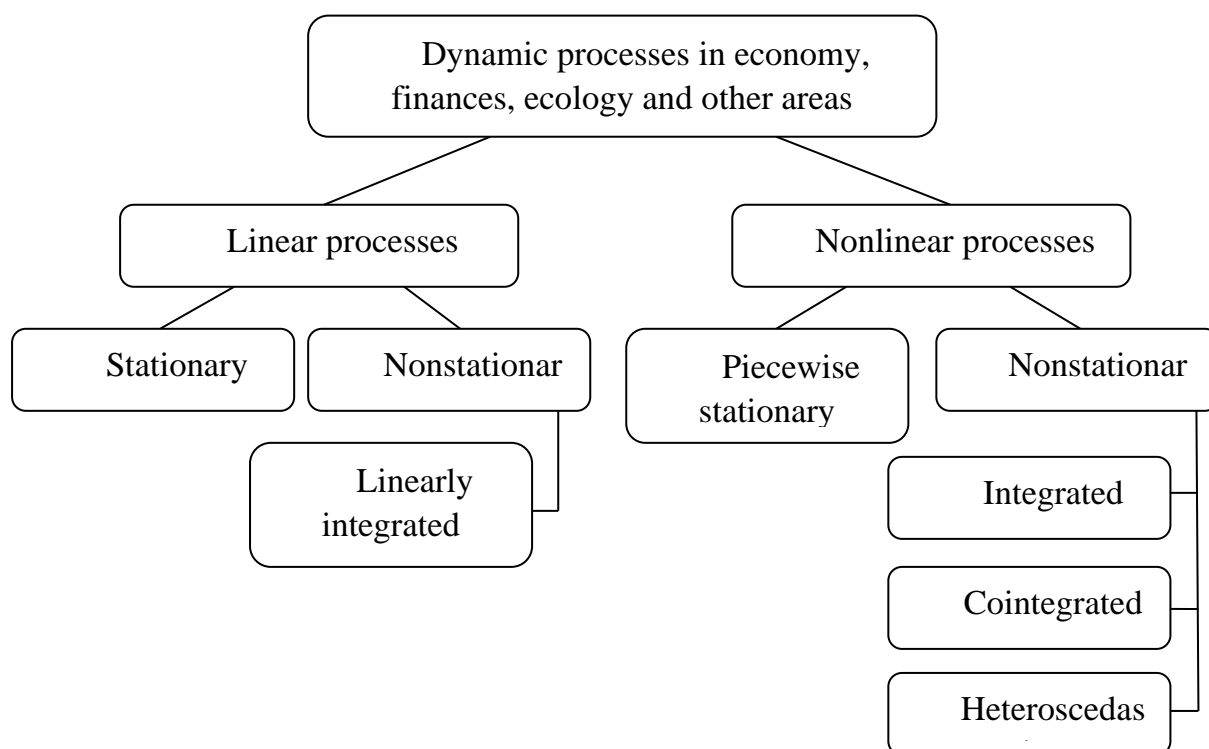


Рисунок 2.1 - Спрощена класифікація динамічних процесів

Лінійні процеси можуть бути стаціонарними без тренда і нестаціонарними, коли вони містять лінійний (першого порядку) тренд, $I(1)$,

де $I(1)$ означає інтегрований першого порядку. Якщо дисперсія (коваріація) стохастичного лінійного процесу залежить від часу, то вона класифікується як гетероскедастична і вимагає нелінійних моделей для опису дисперсії процесу і, можливо, самого процесу.

Існує також велика різноманітність нелінійних процесів, хоча ми вибрали тільки деякі з них, які частіше зустрічаються

2.2.2 Методика моделювання нелінійних нестационарних процесів

Методологія, запропонована для моделювання нелінійних нестационарних процесів, показана на рис. 2.2. На першому етапі методології зібрані дані піддаються попередній обробці, яка може включати в себе наступні основні операції: зобов'язання пропущених спостережень, нормалізацію в заданому діапазоні, цифрову або оптимальну фільтрацію в залежності від постановки задачі, аналіз основних компонентів, відповідну обробку викидів та ін..

Тут же доречно виконати ідентифікацію і усунення (зменшення) невизначеностей даних, які можуть стосуватися наступного: оцінка незмірної величини; обчислення загальних статистичних параметрів (дисперсії, коваріації, середніх, медіани і т.д.); виконання структуризації даних по постановці завдання; аналіз типів розподілу і їх параметрів; оцінка, при необхідності, попередніх ймовірностей. На рис. 2.2 приведено запропоновану функціональну схему системи прогнозування:

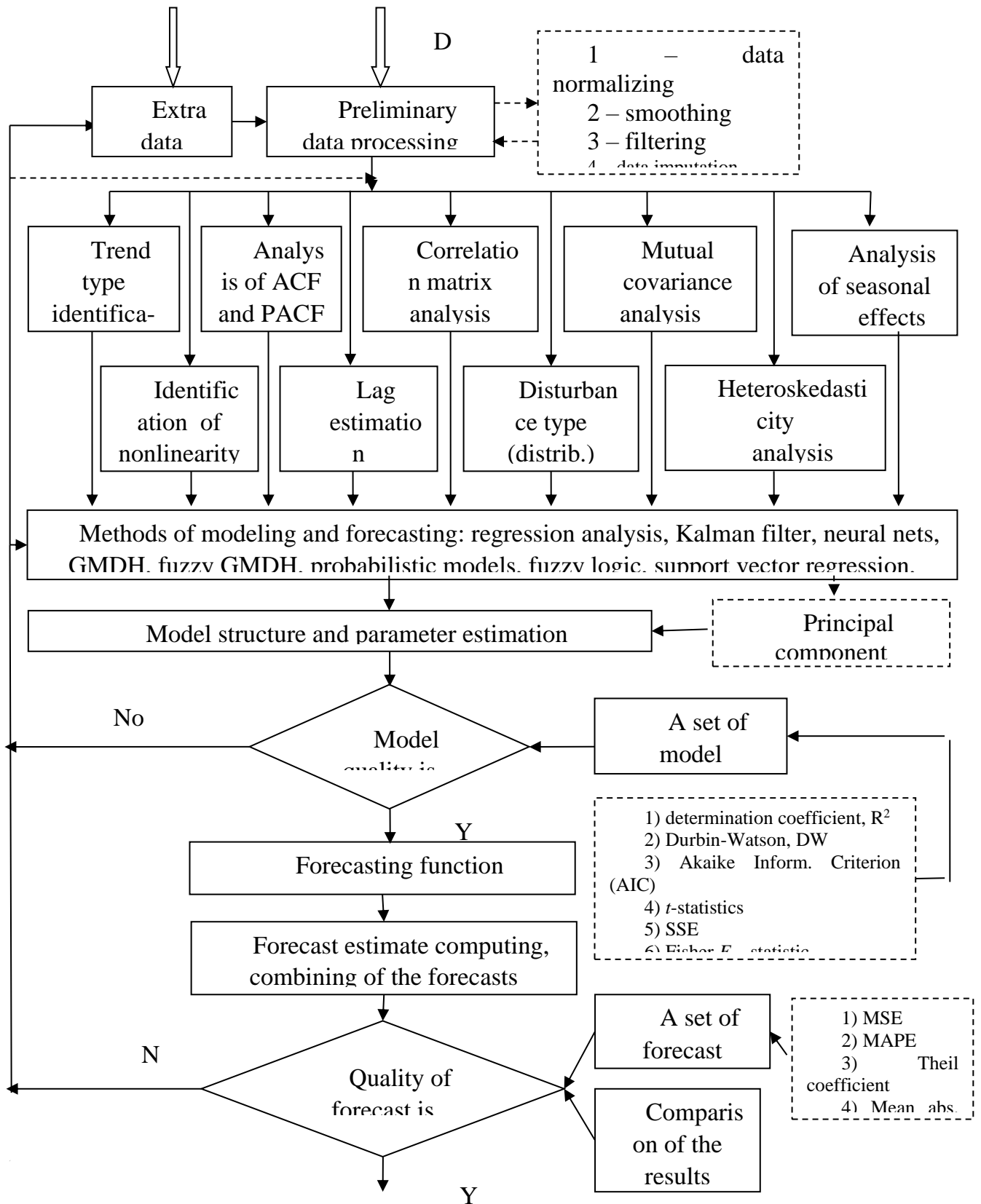


Рисунок 2.2 - Запропонована функціональна схема системи прогнозування

Оцінка структури моделі з використанням статистичного та імовірнісного (взаємного) інформаційного аналізу, що дає можливість оцінити такі елементи структури моделі: розмірність моделі - кількість рівнянь, що створюють модель; порядок моделі (вищий порядок різниці або диференціального рівняння моделі); нелінійність і її тип; оцінка часу затримки входу і тип імовірнісного розподілу для змінних моделі. Для кількох моделей-кандидатів завжди доцільно проводити оцінку структури, щоб мати можливість вибрати кращого з оцінюваних кандидатів.

Формально, для виявлення нелінійності в статистичних даних повинні застосовуватися наявні статистичні тести і методики. На рис. 2.3 показані деякі відомі методи перевірки даних на нелінійність.

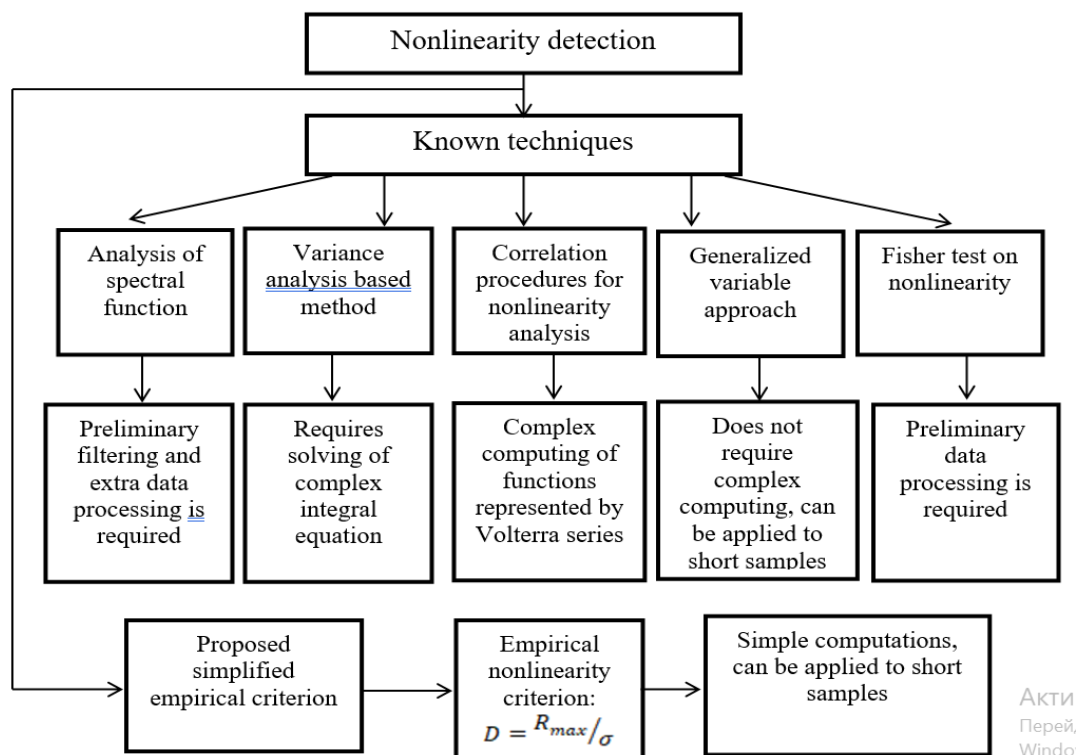


Рисунок 2.3 - Деякі методи тестування даних на нелінійність

Поряд із застосуванням відомої техніки ми запропонували новий спрощений емпіричний критерій для виявлення нелінійності в даних, який показаний нижче на рис. 2.3: тут представлено максимальне відхилення

досліджуваного процесу від його лінійної апроксимації; є зразком стандартного відхилення процесу. Він не вимагає складних обчислень, однак надає додаткову інформацію про наявність нелінійності.

Послідовність операцій, що дозволяє побудувати нелінійну модель, показана на рис. 2.4, фактично це частина загальної процедури побудови моделі, наведеної на рис. 2.2.

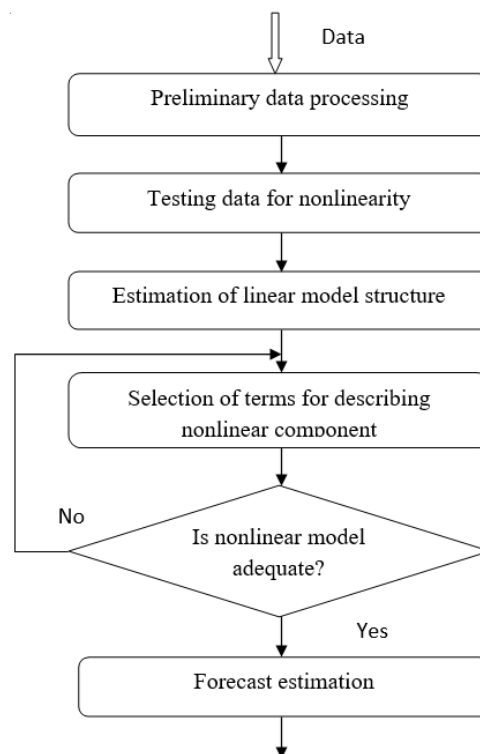


Рисунок 2.4 - Процедура, яка ілюструє пошук формального опису нелінійного процесу

Розглянути можливість опису нелінійностей щодо змінних моделі. Нелінійності можна визначити наступним чином: спочатку оцінюється лінійна частина моделі з використанням лінійних структур, таких як авторегресійні рівняння з ковзним середнім (ARMA (p, q)), можливо, з лінійним трендом, кратною регресією і ін. Потім до лінійної частини додається нелінійна частина, використовуючи такі можливості: опис нелінійного тренда, білінійні або вищі члени. Іноді додаються нелінійні члени, що описують циклічні зміни основної

змінної і т.д. Практика побудови моделей показує, що адекватність моделі може бути досягнута при використанні комбінації лінійної і нелінійної регресії, лінійної регресії і Байєсових мереж, лінійної регресії і нелінійних функцій у вигляді непараметричних ядер. Відповідно до цього підходу можна побудувати кілька моделей-кандидатів з подальшим вибором кращого на основі відповідного набору критеріїв статистичної адекватності, як показано на рис. 2.2. Зрозуміло, що формальні можливості визначення типу нелінійності унікальним способом не завжди існують, наприклад, при коротких вибірках даних.

Наступним кроком є оцінка параметрів моделі з використанням альтернативних методик; в лінійному випадку це звичайні найменші квадрати (OLS) та їх клони, максимальна ймовірність (ML) і багато інших. У разі нелінійної оцінки моделі корисні такі методи: ML, процедури Markov Chain Monte Carlo (MCMC), нелінійні найменші квадрати (NLS) та інші відповідні підходи, здатні забезпечити незсунені оцінки параметрів при конкретних імовірнісних розподілах змінних моделі і структур моделі.

Правильне застосування альтернативних методик оцінки параметрів забезпечує можливість подальшого порівняння моделей-кандидатів і вибору кращої з них. Можна також простежити причини існуючих параметричних невизначеностей в такій формі: оцінки параметрів, обчислені на основі статистичних даних, не можуть бути послідовними, можуть містити зміщення і можуть бути неефективними. Всі ці ефекти в кінцевому підсумку призводять до поганої адекватності побудованої моделі.

На наступному етапі обчислюється набір статистичних параметрів, що характеризують якість моделі (адекватність) і вибір найбільш відповідної моделі з набору моделей-кандидатів. Немає необхідності залишати тільки одну модель для розрахунку прогнозів (або розв'язок контрольного завдання). Знову ж таки, це може бути набір "кращих" моделей, побудованих за різними ідеологіями. Остаточний вибір завжди робиться після застосування моделей для вирішення задачі у відповідності з початковою постановкою завдання.

Після розрахунку прогнозів процесу (який досліджуємо) з використанням моделей-кандидатів для вибору кращого результату застосовується інший набір критеріїв якості прогнозу, наприклад, середня абсолютна процентна помилка (MAPE), коефіцієнт Тейла, середня абсолютна помилка (MAE), мінімальна і максимальна помилки прогнозування та ін. Побудовані моделі також повинні бути протестовані на аналогічному процесі, тобто виконане калібрування моделі.

На цьому етапі можна зробити висновок, що наявність зазначених невизначеностей даних, а також необхідність ієрархічної побудови системи обробки даних з особливостями адаптації та оптимізації (структурної та параметричної) вимагають застосування сучасного системного підходу, що забезпечує можливість успішного та якісного вирішення проблем, що виникають при статистичній обробці даних, побудові математичної моделі, прогнозної оцінки і генерації альтернатив рішень.

У даній роботі запропоновані деякі практичні можливості побудови процедур обробки даних на основі сучасних принципів системного підходу.

2.2.3 Робота з невизначеностями

Важливим моментом, який необхідно враховувати при побудові моделей, оцінці прогнозів і прийнятті рішень, є проблема усунення невизначеностей. Невизначеності розглядаються, як фактори негативного впливу на всю процедуру обробки даних і побудови моделі, які можуть привести до незадовільних проміжних і кінцевих результатів обчислювальних експериментів. Ці фактори неминуче виникають при використанні статистичних даних для побудови моделі.

Джерелами можливих невизначеностей при побудові моделі і оцінки прогнозу є наступні:

- деякі дані недоступні або втрачені з різних причин, що вимагає застосування відповідних процедур поставлення даних;
- дані генеруються досліджуваною системою, на яку впливають випадкові зовнішні обурення, що спотворюють фактичні значення стану системи;
- спостереження завжди вимірюються (збираються) з деякими похибками, вплив яких має бути зведено до мінімуму перед подальшим використанням вимірювань;
- параметрами, що визначають структуру моделі, також присвоюються випадкові значення, що призводить до приблизної структури моделі;
- дуже часто виникають труднощі з вибором методу оцінки параметрів моделі, особливо у випадках коротких вибірок або вибірок з промахами, або коли розподіл ймовірностей даних погано визначено; в результаті оцінки параметрів можуть бути упередженими або неефективними;
- багатоступінчате прогнозування вимагає використання проміжних оцінок, що може привести до істотного погіршення прогнозних оцінок.

Таким чином, процедури побудови моделей, які зазвичай реалізуються в рамках відповідним чином спроектованих систем підтримки прийняття рішень, повинні містити засоби для ідентифікації невизначеностей і мінімізації їх негативного впливу. У табл. 2.1 наведені типи можливих невизначеностей і відображені деякі можливі засоби їх усунення.

Таблиця 2.1 - Можливі типи невизначеностей при моделюванні і прогнозуванні

Но.	Тип невизначеності	Причина невизначеності	Методи мінімізації впливу невизначеності
1	Невизначеність структури моделі	<ul style="list-style-type: none"> – Неможливість встановлення усіх причинно-наслідкових зв'язків між змінними – Приблизні значення для параметрів структури моделі 	<ul style="list-style-type: none"> – Експертний підхід – Застосування статистичних тестів – Застосування теорії тестування гіпотез

2	Статистична невизначеність	<ul style="list-style-type: none"> – Похибка вимірювання – Стохастична деформація – Відхилення – Відсутні значення даних; 	<ul style="list-style-type: none"> – Цифрові та оптимальні фільтри – Уточнення типу розподілу – Теорія екстремальних значень – Вставка значень, яких не вистачає
3	Параметрична невизначеність	<ul style="list-style-type: none"> – Неправильний вибір метода оцінки параметрів – Короткі приклади 	<ul style="list-style-type: none"> – Застосування альтернативних методів оцінки параметрів – Розширення зразків даних;
4	Ймовірнісна невизначеність	<ul style="list-style-type: none"> – Складні механізми причинно-наслідкових зв'язків між змінними 	<ul style="list-style-type: none"> – Байєсівські мережі – Моделі Маркова – Ймовірнісні фільтри – Умовні дистрибуції
5	Невизначеність амплітуди	<ul style="list-style-type: none"> – Змінні, які не можна виміряти – Великі похибки вимірювання 	<ul style="list-style-type: none"> – Bayesian data processing; – Нечітка логіка.

Робота з невизначеностями структури моделі.

При використанні DSS структура моделі практично завжди повинна оцінюватися з використанням даних. Це означає, що елементи структури моделі практично завжди беруть тільки приблизні значення. При побудові моделі для прогнозування ми будемо кілька кандидатів і вибираємо кращого з них, використовуючи набір статистики якості (адекватності) моделі. В цілому для боротьби зі структурними невизначеностями можна виділити наступні методи:

- поступове уточнення порядку моделі (для структур AR (p) або ARMA (p, q)) із застосуванням адаптивного підходу до моделювання та автоматичного пошуку "кращої" структури з використанням складних критеріїв якості статистики;
- адаптивна оцінка часу затримки вхідних даних (лага) і типу розподілу даних з її параметрами;

- формальний опис виявлених нелінійностей процесу з використанням альтернативних аналітичних форм з подальшою оцінкою адекватності моделі і якості прогнозу.

Простий приклад складної моделі і критерії прогнозування може виглядати наступним чином:

$$J = |1 - R^2| + |2 - DW| + \beta \ln(MAPE) \rightarrow \min_{\hat{\theta}_i} \quad (2.1)$$

Або в більш складній формі:

$$J = |1 - R^2| + \alpha \ln \left[\sum_{k=1}^N e^2(k) \right] + |2 - DW| + \beta \ln(MAPE) + U \rightarrow \min_{\hat{\theta}_i} \quad (2.2)$$

де R^2 - коефіцієнт детермінації;

$\sum_{k=1}^N e^2(k) = \sum_{k=1}^N [y(k) - \hat{y}(k)]^2$ - сума квадратів похибок моделі;

DW - статистика Дарбіна-Уотсона;

$MAPE$ - середня абсолютна похибка у відсотках для однокрокового прогнозу;

U - коефіцієнт Тейла, який характеризує якість побудованого прогнозу моделі;

α, β - правильно підібрані вагові коефіцієнти;

$\hat{\theta}_i$ - вектор параметрів для $(i - th)$ моделі-кандидата.

Комбінований критерій цього типу використовується для автоматичного відбору побудованої моделі кращого кандидата. Представлені критерії також дозволяють працювати з DSS в адаптивному режимі. Очевидно, що інші форми комбінованого критерію можливі в залежності від конкретної мети

побудови моделі. Що важливо при побудові критерію: чи не переважають окремі члени в правій частині, що пригнічувало б інші компоненти.

Усунення невизначеностей типу рівня (амплітуди).

Наявність випадкових і/або невимірних змінних призводить до необхідності використання нечітких множин для опису процесів в таких ситуаціях. Змінна з випадковою амплітудою може бути описана з деяким розподілом ймовірностей, якщо вимірювання доступні або коли вони надходять на аналіз в прийнятний проміжок часу. Однак деякі змінні не можуть бути виміряні в принципі, наприклад, сума тіньового капіталу, який "зникає" щомісяця в офшорних зонах, або сума тіньових зарплат, що виплачуються на якійсь компанії, або технологічний параметр, який не може бути виміряний в режимі реального часу через відсутність відповідного датчика або фізичних труднощів на місці. У таких ситуаціях змінній можна привласнити набір характерних значень в лінгвістичній формі, скажімо так: сума капіталу = {дуже низький, низький, середній, високий, дуже високий}. Для таких нечітких змінних існує повний набір необхідних математичних операцій. Нарешті, нечітка величина може бути перетворена в точну нечітку форму за допомогою відомих методів трансформації.

Ймовірнісні невизначеності і їх опис.

Використання випадкових величин призводить до необхідності оцінки фактичних розподілів ймовірностей і їх застосування в обчислювальних процедурах виведення. Зазвичай значення, яке спостерігають, відоме тільки приблизно, хоча відомі межі фактичних значень. Відповідні розподіли ймовірностей дуже корисні для опису досліджуваних процесів в таких ситуаціях. При роботі з дискретними наслідками ми присвоюємо ймовірності конкретним наслідкам за допомогою функції мас. Вона показує, скільки "ваги" (або маси) слід присвоїти кожному спостереженню або вимірюванню. Відповіддю на запитання про значення конкретного результату буде його

маса. Аксиоми ймовірностей Колмагорова корисні для більш глибокого розуміння того, що відбувається. Якщо одночасно аналізуються дві і більше змінні, то необхідно побудувати і використовувати спільні розподіли. Спільні розподіли дозволяють при необхідності оцінювати умовні ймовірності за допомогою процедур перенормування.

Дуже корисним для виконання імовірнісних обчислень є поняття умовної незалежності:

$$P(x, y | z) = P(x | z) P(y | z), \quad (2.3)$$

де x та y - незалежні події.

Такі ідентичності дуже зручні, хоч треба й бути обережним при їх використанні, тобто події повинні бути дійсно незалежними. Чудовий інтуїтивний сенс дискретного закону Байєса

$$P(A/B) = P(B/A) P(A) / P(B) \quad (2.4)$$

який дозволяє ставити обернені питання: «З урахуванням того, що подія A відбулася, яка ймовірність того, що конкретна подія B викликала його?» Гранична ймовірність $P(B)$, може бути розрахована з використанням відповідних умов. Ймовірність того, що подія B відбудеться в цілому, $P(B)$, можна отримати при наступній умові:

$$P(B) = P(B/A) P(A) + P(B/\bar{A}) P(\bar{A}) \quad (2.5)$$

Ймовірнісні типи невизначеностей щодо того, відбудеться чи ні яка-небудь подія, можуть бути взяті до уваги при використанні імовірнісних моделей. Для вирішення завдання опису і обліку таких невизначеностей можуть бути взяті різні байєсовські моделі, які розглядаються як байєсовський програмуючий формалізм. Набір моделей включає байєсовські мережі (BN), динамічні байєсовські мережі (DBN), байєсовські фільтри, фільтри частинок, приховані марківські моделі, калмановські фільтри, байєсовські карти і ін. Структура Байєсової програми включає в себе наступні елементи:

(1) опис задачі та формулювання постановки задачі з основним питанням форми:

$$P(\textit{Searched} / \textit{Known}) \text{ або } P(X_i / D, Kn) \quad (2.6)$$

де X_i визначає одну змінну;

(2) використання попередніх знань Kn та експериментальних даних D для визначення структури моделі і параметрів;

(3) вибір і застосування відповідної техніки формування висновків для відповіді на раніше поставлене питання;

(4) перевірка якості кінцевого результату. Такий підхід також добре працює в режимі адаптації з ціллю адаптації структури і параметрів моделі, яка розробляється, до нових експериментальних даних або до нового режиму роботи системи, наприклад, для оцінки попередніх розподілів або структури BN.

Тоді завдяки наступним принципам системного аналізу, а саме:

- принцип системної координації функцій;
- принцип процесуальної повноти;
- принцип функціональної ортогональності;
- принцип залежності взаємної інформації між реалізованими функціями;

- принцип функціональної раціональності;
- принцип багатоцільового узагальнення;
- принцип багатфакторної адаптації,
- принцип раціонального доповнення

можна реалізувати спеціалізовану ДСС для моделювання і прогнозування, всі функціональні модулі якої розроблятимуться з необхідним ступенем узагальнення, що забезпечить можливість досягнення високоякісних результатів вирішення безлічі можливих завдань, що відносяться до одного класу.

Це може бути якісне прогнозування та генерація альтернативи рішення щодо майбутньої еволюції лінійних або нелінійних нестационарних процесів. Серед яких можна виділити наступні: накопичення необхідних даних і їх попередня обробка; оцінка структури і параметрів для набору математичних моделей-кандидатів; побудова прогнозних функцій на розроблених моделях і обчислення відповідних прогнозів; відбір найкращих результатів обчислень з використанням відповідних наборів критеріїв якості.

2.2.4 Модель авторегресії (АР)

Будь-який часовий ряд можна описати за допомогою авторегресійної моделі, рівняння якої має наступний вигляд:

$$y(k) = a_0 + \sum_{i=1}^n a_i y(k-i) + \varepsilon(k) \quad (2.7)$$

Особливістю авторегресійних моделей є те, що у правій частині рівняння містяться значення минулих вимірів, отже значення часового ряду в

конкретний момент часу лінійно залежить від значень у попередніх моментах часу.

Дана модель є найпростішою в своєму класі, і часто вона недостатньо адекватно може описувати складні процеси, проте часто використовується для моделювання сезонності. Попри свою простоту, вона також є основою для більш складних моделей, які будуть розглянуті далі.

2.2.5 Авторегресія з ковзним середнім порядком (p, q) (АРКС)

В загальному вигляді авто регресійне рівняння має вигляд:

$$y(k) = a_0 + \sum_{i=1}^p a_i y(k-i) + \sum_{j=1}^q b_j \varepsilon(k-j) + \varepsilon(k) \quad (2.8)$$

де $a_0 + \sum_{i=1}^p a_i y(k-i)$ - авторегресійна складова,

$\sum_{j=1}^q b_j \varepsilon(k-j)$ - ковзне середнє,

$\varepsilon(k)$ - некорельована випадкова величина.

АРКС (p, q) можна побудувати по двом різним методикам:

Варіант А

Побудова АРКС (p, q), коли ковзне середнє будується по залишкам АР (p) рівняння моделі:

1). Оцінка коефіцієнтів рівняння моделі АР (p):

$$y(k) = a_0 + \sum_{i=1}^p a_i \cdot y(k-i) + e \quad (2.8)$$

де e - залишки моделі (*resid*).

2). По залишкам АР (p) формують КС.

3). Визначення порядку КС.

Будується ЧАКФ вектора для зберігання ковзного середнього та визначають q.

4). Оцінка коефіцієнтів КС(q).

Визначаються коефіцієнти рівняння

$$resid(k) = mv(k) + \sum_{j=1}^q b_j \cdot mv(k-j) \quad (2.9)$$

Побудова АРКС (p, q) здійснюється по частинам, використовуючи рівняння залишків та АР (p). Звичайно можна оцінити всі коефіцієнти рівняння одразу, проте вважається, що оцінка початку коефіцієнтів авторегресійної частини, а потім коефіцієнтів при ковзних середніх дає більш якісний результат.

$$y(k) = a_0 + \sum_{i=1}^p a_i \cdot y(k-i) + mv(k) + \sum_{j=1}^q b_j \cdot mv(k-j) \quad (2.10)$$

Варіант В

Побудова АРКС (p, q), коли ковзне середнє будується по вихідному сигналу у:

1). Побудова КС по вихідному сигналу у, створюємо вектор mv .

2). Визначення порядку КС (q).

Будується ЧАКФ вектора для зберігання ковзного середнього та визначають q. Зазвичай, коли КС будують по вихідному сигналу у, дуже часто $p=q$.

3). Оцінка усіх коефіцієнтів рівняння АРКС (p, q).

При цьому, здійснюється перенесення $mv(k)$ в ліву частину рівняння для покращення якості отриманих коефіцієнтів.

$$y1(k) = y(k) - mvk = a_0 + \sum_{i=1}^p a_i \cdot y(k-i) + \sum_{j=1}^q b_j \cdot mv(k-j) \quad (2.11)$$

Якщо вірити теорії, коефіцієнти мають вийти: $\sum_{i=1}^p a_i < 1$, а $\sum_{j=1}^q b_j \rightarrow 1$.

На практиці так виходить не завжди і для кожного окремого випадку визначити погано це чи добре можна, проаналізувавши результати прогнозування за допомогою отриманої моделі на декілька років вперед. Тобто при отриманні задовільних результатів невиконання цієї умови прийнятне.

2.2.6 Загальний розв'язок SIR-модель

Динаміка розвитку епідемії, наприклад, грипу, часто набагато швидша, ніж динаміка народження і смерті населення, тому, в простих полігамних моделях життєвий цикл населення досить часто опускається. Методи такого типу були застосовані для дослідження певної дитячої хвороби в Чернівцях в 1988-1989 роках.

SIR модель без врахування життєвого циклу населення(народження і смерті, або інколи називають демографією) може бути описана наступною системою диференціальних рівнянь:

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta IS}{N}, \\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I. \end{aligned} \quad (2.12)$$

Ця система є нелінійною й немає узагальненого аналітичного розв'язку. Проте, певні результати цієї моделі можуть бути отриманими аналітично.

По-перше,

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0, \quad (2.13)$$

звідси $S(t) + I(t) + R(t) = \text{Constant} = N$,

де N - чисельність населення, яка вважається сталою.

По-друге, слід зазначити, що динаміка інфекційного захворювання залежить від коефіцієнта поширення інфекції R_0 :

$$R_0 = \frac{\beta}{\gamma}, \quad (2.14)$$

Це співвідношення показує число нових випадків поширення інфекції, де всі особи є сприйнятливими до захворювання.

Позначимо час контакту між особою й $T_c = \beta^{-1}$ інфекцією $T_r = \gamma^{-1}$ а час одужання. Звідси випливає, що, в середньому, кількість контактів зараженої T_r/T_c людини з іншими людьми, перш ніж вона одужає дорівнюватиме

$$S(t) = S(0)e^{-R_0(R(t)-R(0))/N} \quad (2.15)$$

де $S(0)$ і $R(0)$ початкові значення сприйнятливих до захворювання й тих, що одужали осіб відповідно. Таким чином спрямувавши $t \rightarrow +\infty$, частка осіб, які одужали відповідатиме наступному рівнянню:

$$R_\infty = N - S(0)e^{-R_0(R_\infty-R(0))/N} \quad (2.16)$$

Це рівняння показує, що в кінці епідемії, навіть якщо $S(0) = 0$, не всі особи популяції одужали, тому повинно залишатись певне число осіб сприйнятливих до захворювання. Це означає, що кінець епідемії спричинений скороченням числа інфікованих людей, а не повною відсутністю сприйнятливих до захворювання осіб. Роль коефіцієнта поширення інфекції є надзвичайно важливою. Справді, переписавши рівняння зміни кількості інфікованих осіб наступним чином:

$$\frac{dI}{dt} = (R_0 S/N - 1)\gamma I, \quad (2.17)$$

Тоді за умов $R_0 > \frac{N}{S(0)}$, отримаємо $\frac{dI}{dt}(0) > 0$,

тобто відбудеться спалах епідемії зі зростанням числа інфікованих осіб.

І навпаки, $R_0 < \frac{N}{S(0)}$ якщо, тоді $\frac{dI}{dt}(0) < 0$,

тобто незалежно від початкової кількості сприйнятливих до захворювання осіб, хвороба ніколи не зможе спричинити спалах епідемії.

2.3 Інструменти для моделювання і прогнозування обраних процесів

Основною проблемою предметної області є велика кількість даних та критеріїв аналізу даних для адекватного прогнозування. Тому необхідне застосування комп'ютерної техніки для аналізу даних про захворювання. І при цьому більшість систем імітаційного моделювання, що використовуються сьогодні, не пристосовані до таких завдань і не дають змоги досягти високої якості прогнозування. Наприклад:

1. EpiGrass – це програмний засіб для мереж епідеміологічного моделювання та аналізу. Охоплює епідеміологічні дані та моделі передавання хвороб і контролю. EpiGrass розрахований на побудову та моделювання великомасштабних популяційних моделей.

EpiGrass написана мовою програмування Python, яка значною мірою сприяє гнучкості всієї системи в зв'язку з динамічним характером мови програмування. Епідеміологічні процеси, що мають географічну прив'язку мережі, можуть бути подані в об'єктно-орієнтованому вигляді.

Система EpiGrass має графічний користувацький інтерфейс (GUI), який обробляє декілька вхідних файлів, які потрібні для визначення моделей, виконує симуляцію та генерує результуючу інформацію.

Результат EpiGrass подає різними способами. 3D карта з анімацією доступна прямо в GUI. Інші вихідні дані можуть бути подані через ГІС програми зі згенерованого файла формату shapefile. KML файл також може бути переглянутий в Google Earth.

2. Model-Builder – це графічна утиліта для дизайну, симуляції та аналізу математичних моделей на основі диференціальних рівнянь. Як і EpiGrass, Model-Builder не є ГІС.

Model-Builder дає змогу зберігати результати у найпоширеніших графічних форматах: PNG, SVG, PDF тощо. До речі, з електронних таблиць результатів можна зробити індивідуальні ділянки зі змінним. Підтримується експортування даних у текстовий файл формату *.csv. Безкоштовне програмне забезпечення (Ліцензія GPL) надає латекс-обробку системи рівнянь та інтуїтивний графічний інтерфейс. Це багатоплатформенна програма, яка запускається там, де працює Python.

3. AnyLogic – це інструмент імітаційного моделювання, що об'єднав методи системної динаміки, "процесного" дискретно-дійового й агентного моделювання в одній мові і одному середовищі розробки моделей. Гнучкість

AnyLogic дає змогу відображати динаміку складних і різномірних економічних і соціальних систем на будь-якому бажаному рівні абстракції.

AnyLogic містить набір примітивів і бібліотечних об'єктів для ефективного моделювання виробництва і логістики, бізнес-процесів і персоналу, охорони здоров'я, фінансів, споживчого ринку, а також 159 навколишньої інфраструктури в їх природній взаємодії. Об'єктно-орієнтований підхід, запропонований AnyLogic, полегшує ітеративну та поетапну побудову великих моделей.

AnyLogic 6 заснований на Java на платформі Eclipse – сучасному стандарті для бізнес-додатків. Завдяки Eclipse AnyLogic працює на всіх поширених операційних системах (Windows, Mac, Linux тощо).

4. Echo є моделювальним інструментом з відкритим вихідним кодом, який розроблений для вивчення механізмів, які регулюють опрацювання інформації в системах, у які входить багато взаємодіючих адаптивних агентів. З Echo можна моделювати моделі екосистем, в якій розвивається агенти, що перебувають в умовах обмежених ресурсів навколишнього середовища. Він також дозволяє індивідуальні генотипні правила кодування для взаємодії. Echo працює на всіх UNIX та LINUX системах.

5. (STEM) - (переклад. "просторово-часовий епідеміологічний симулятор") є вільним програмним забезпеченням Eclipse Foundation, фреймворком та розробницьким інструментом, покликаним допомогти вченим у створенні та використанні просторових і часових моделей інфекційних хвороб. Архітектура також підтримує те, що користувачі можуть не тільки створювати нові моделі і скласти новий сценарій, але і обмінюватися цими моделями і сценаріями як компонентами, що багато разів використовуються, і тим самим впливати на роботу один одного. Це проект з відкритим кодом.

6. EViews – це статистичний пакет для Windows, який пропонує академічним дослідникам, корпораціям, державним установам та студентам доступ до потужних інструментів статистики, прогнозування та моделювання, дає можливість проведення регресійного аналізу через інноваційний, простий у користуванні об'єктно-орієнтований інтерфейс. Використовується в основному для часових рядів.

Графічні можливості пакета Eviews, незважаючи на свою простоту, забезпечують основні формати представлення даних необхідні для успішної роботи аналітика.

У даній дипломній роботі будуть використовуватися для моделювання та прогнозування процесів поширення коронавірусу AnyLogic, EViews.

РОЗДІЛ 3: ПОБУДОВА МОДЕЛЕЙ ТА ОЦІНЮВАННЯ ПРОГНОЗІВ ПРОЦЕСІВ ПОШИРЕННЯ ПАНДЕМІЇ

3.1 Загальний підхід до моделювання ситуації поширення пандемії вірусу

На рис. 3.1 представлено графік динаміки росту числа тих, хто захворів за межами Китаю в період від 22 січня до 6 березня 2020 року:

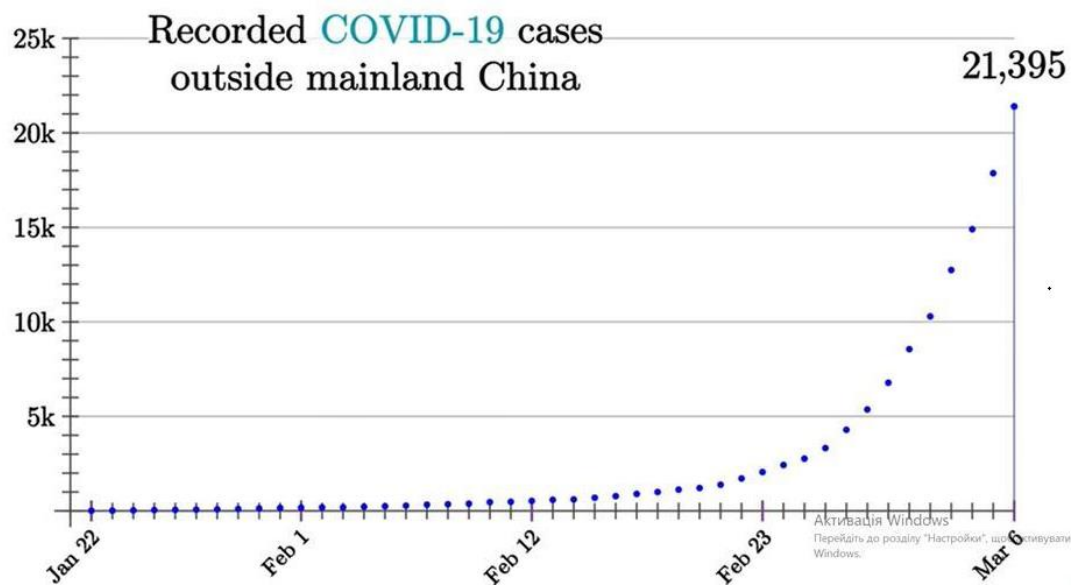


Рисунок 3.1 Кількість захворівших за межами Китаю

Спостерігається яскравий експоненціальний ріст, тобто кожного наступного дня кількість заражених осіб множить на певну константу, яка у випадку з коронавірусом варіює в межах 1,15 – 1,25. Для вірусів властива така природа росту тому, що чим більше інфікованих особин, тим швидше поширюється сам вірус в популяції.

Наприклад, нехай кожна інфікована вірусом людина в середньому контактує з E людей в день. Введемо ймовірність зараження вірусом p для кожного контакту. Тоді кількість нових випадків зараження кожного дня можна порахувати за формулою:

$$\Delta N_d = E \cdot p \cdot N_d \quad (3.1)$$

де N_d – кількість випадків в день,

E – середня кількість людей, які контактують з хворим в цей день,

p – ймовірність зараження при контакті.

Той факт, що N_d з кожним днем росте, суттєво прискорює ріст інфікованих, що й спостерігається на графіку вище. Тож додаючи кожного дня ці нові випадки, отримуємо:

$$\begin{aligned} N_{d+1} &= N_d + E \cdot p \cdot N_d \\ N_{d+1} &= (1 + E \cdot p) \cdot N_d \end{aligned} \quad (3.2)$$

Зрештою, це те саме, що помножити на деяку константу, більшу за 1. Як згадувалося, раніше у випадку з коронавірусом її значення варіюється в межах 1,15 – 1,25. Для кращого розуміння серйозності ситуації та швидкості поширення вірусу в популяції переведемо вісь ординат в логарифмічну шкалу (див. рис. 3.2):

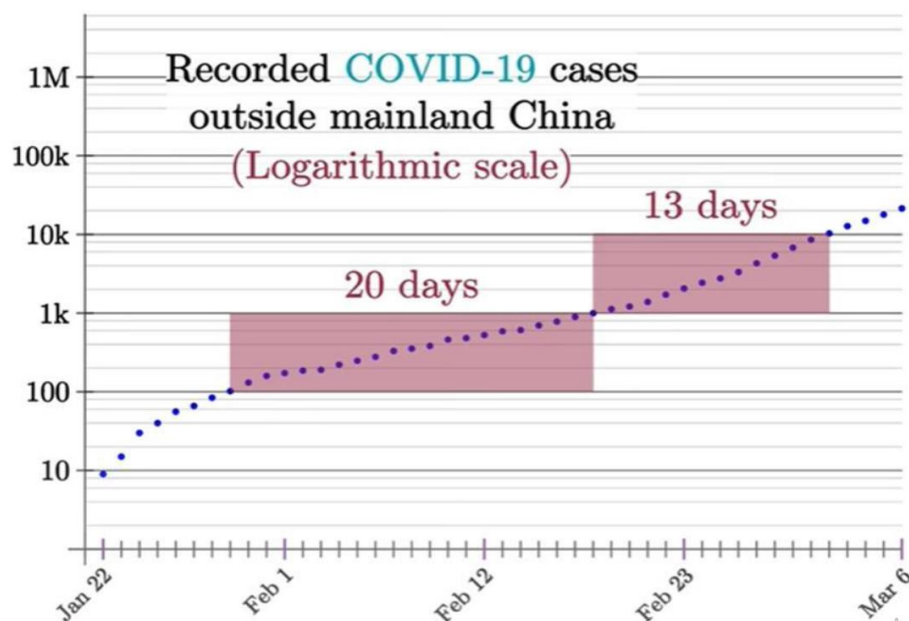


Рисунок 3.2 Кількість захворівших на логарифмічній шкалі

В даному масштабі експоненціальний ріст виглядає майже, як пряма лінія. В нашому випадку виявилось достатнім 20 днів аби перейти від 100 до 1000 випадків і 13 днів аби перейти від 1000 до 10 000 випадків. Якщо за допомогою методу лінійної регресії визначити пряму залежності (експоненціальна апроксимація $R^2 = 0,975$), то можна з нахилу прямої зауважити, що в середньому наш графік зростає в 10 разів кожні 16 днів (див. рис. 3.3):

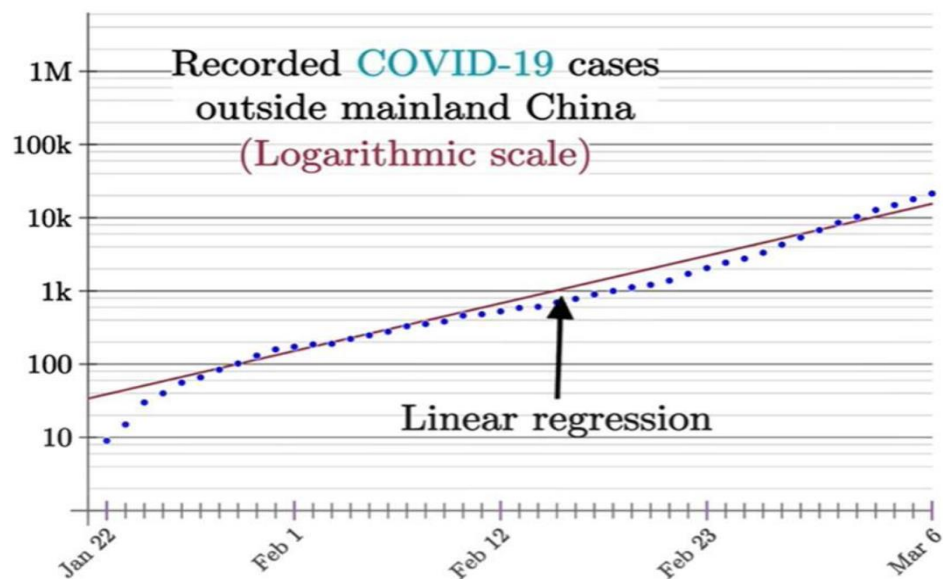


Рисунок 3.3 Графік з визначеною прямою залежності

Виходить, коли в реальності переглядаєш статистику захворюваності в світі і в одній країні 6000 випадків, а в другій тільки 60 – не правильним буде сказати, що друга країна справляється в 100 разів краще, ніж перша. Просто зважаючи на отримані результати апроксимації та збільшення числа заражених в 10 разів в середньому кожні 16 днів, можна сказати, що в першу країну вірус попав приблизно на місяць раніше.

Такими темпами вже через:

- 30 днів за підрахунками по нашій моделі число інфікованих дорівнює 1 млн.

(насправді станом на 5 квітня світова статистика говорила про 1 202 827 осіб, з яких на той момент уже померло 64 774).

- 47 днів це число має досягти відмітки 10 млн.

(в реальності станом на 22 квітня число заражених по світу склало всього 2 567 495 осіб, з яких тоді померло 177 935)

- 64 дні отримуємо результат рівний 100 млн.

інфікованих.

(реальна картина на 9 травня склала 4 023 969 інфікованих та 279 311 померло)

- 81 день і число заражених мало б досягнути відмітки в 1 млрд. осіб.

(насправді станом на 26 травня 5 590 364 зареєстрованих випадків і 352 093 смерті)

Чому ж така різниця між даними, отриманими в моделі, та реальною ситуацією? Побудована модель є одною з найпростіших, які взагалі можна скласти для елементарного аналізу та оцінки ситуації поширення пандемії коронавірусу і будь-якого іншого інфекційного захворювання. При її розробці абсолютно не враховано вплив кліматичних умов, швидку випадкову мутацію штаму вірусу, умови співіснування та взаємодії між людьми, варіант введення владою постраждалих країн пакету примусових, обмежуючих свободу громадян мір та заходів, направлених на зниження ймовірності зараження та зменшення контактів за день для кожного до мінімуму, і т.д.

І як знати наперед, що новий вірус схожий на спалах атипової пневмонії в 2002 році, коли було зареєстровано 8096 випадків чи, можливо, більше на пандемію іспанки в 1918 році, коли по різних оцінках заразилося близько 513 000 000 осіб (27% населення світу)?

Модель описує ідеальну для вірусу спрощену до примітиву ситуацію зі сталими тенденціями поширення пандемії (епідемії), але навіть в ній колись ріст повинен почати уповільнюватися при скінченній кількості особин в

популяції. Що є логічним, як мінімум тому, що в якийсь момент інфікована людина почне зустрічати все менше здорових (ще не хворівших, або вже з імунітетом) людей навколо і через певний період часу поширення припиниться, бо заражати банально не стане кого. Залишилось тільки визначити цей момент?

Хоча й виходить, що скористатися лінійною регресією і провести лінію через набір даних – це не найкращий спосіб зробити прогноз. Проте, очевидно при виникненні спалаху будь-якого інфекційного захворювання на початку для нього властивий експоненціальний ріст – ситуація, коли число нових випадків захворювання кожного дня пропорційно збільшується відносно показників попередніх днів, тобто множаться на певну константу. У випадку з пандемією коронавірусу цей процес спостерігався до 5 квітня приблизно, про що свідчить співпадіння реальних даних з отриманими в моделі. Звідси отримали:

$$N_d = (1 + E \cdot p)^d \cdot N_0 \quad (3.3)$$

Що дає можливість зробити короткостроковий прогноз на d днів вперед. Однак через певний проміжок часу значення E і p почнуть невпинно зменшуватись. Тому це потрібно врахувати в даній моделі (3.3) - модифікувати ймовірність зараження, увівши такий фактор, який буде враховувати ймовірність контакту з людиною, ще не інфікованою:

$$p = \left(1 - \frac{N_d}{N_w}\right) \quad (3.4)$$

де N_w – населення планети (pop. size).

В результаті отримаємо логістичну криву (див. рисунок 3.4), яка на початку повністю повторює експоненціальну, але при наближенні до загальної чисельності населення вирівнюється, як і сподівалося. Число нових

інфікованих осіб представлено у вигляді нахилу цієї кривої: на початку спостерігається зростання кількості нових випадків кожного дня, потім певний час крива нахилу не змінює, що означає приблизно однакову кількість виявлених випадків щодня, коли графік (див. рисунок 3.4) проходить точку перегину число нових випадків починає зменшуватись до 0, тобто моменту, коли заражати вже немає кого.

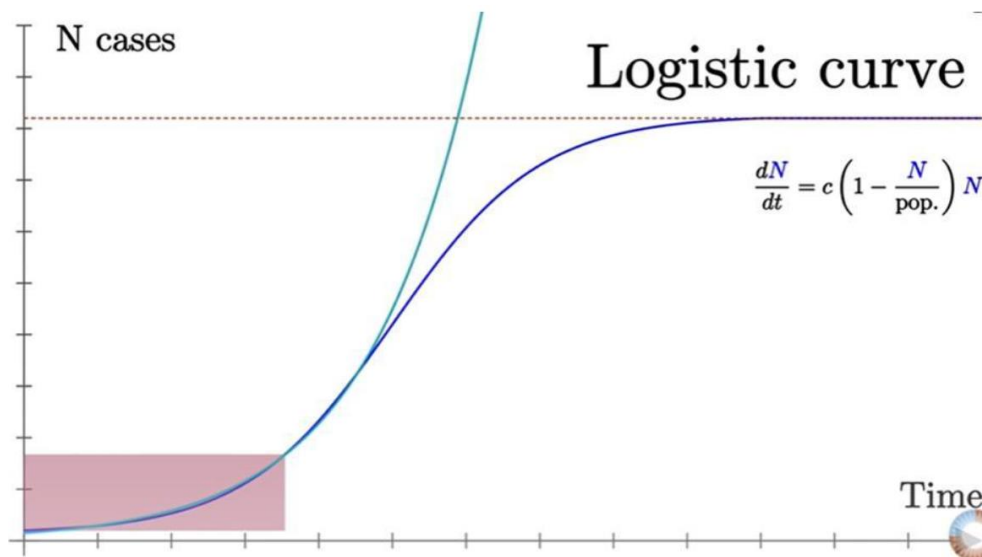


Рисунок 3.4 Графік з урахуванням ймовірності контакту з людьми

Отже, показник за яким доцільно слідкувати під час епідемії (пандемії) це «фактор росту»:

$$\text{Growth factor} = \frac{\Delta N_d}{\Delta N_{d-1}} \quad (3.5)$$

де ΔN_d – кількість нових випадків захворювання,

ΔN_{d-1} – кількість нових випадків зараження за минулий день.

«Фактор росту» (3.5) завжди буде більше 1, поки графік зростає експоненціально, Наближення його значення до 1 повідомить про те, що графік підійшов до точки перегину. При значенні показника рівному 1 графік знаходиться в точці перегину і значить було досягнуто максимального числа нових випадків інфікованих за день, яке надалі буде зменшуватися до нуля.

В реальності на «фактор росту» (3.5) впливає не тільки відсоток інфікованого населення. Рівень зараження (p) знизиться, якщо люди ретельніше і частіше банально митимуть руки, рівень впливу захворювання (E) знизиться, якщо перестати просто збиратися в компанії, подорожувати.

Візьмемо початкову кількість заражених нехай 1000 осіб. При значенні константи $E \cdot p = 0,15$ за 61 день:

$$(1,15)^{61} \cdot 1000 = 5\,041\,598,56 \text{ інфікованих}$$

Якщо зменшити значення хоча б $E \cdot p = 0,05$, за 61 день отримаємо:

$$(1,05)^{61} \cdot 1000 = 19\,613,145 \text{ інфікованих}$$

Зменшивши вплив вірусу в лиці константи $E \cdot p$ всього в 3 рази, кількість можливих заражених за експоненціальний період протікання епідемії зменшиться в 257 разів.

Тобто, якщо люди виявляють занепокоєння з приводу цієї ситуації, то в принципі переживати немає приводу, досить уважніше просто ставитися до власної гігієни і безпеки. Варто почати переживати, якщо більшість ігноруватиме дану ситуацію.

Тому розглянемо можливі дві фундаментальні стратегії боротьби з поширенням пандемії коронавірусу:

а) Пригнічення.

Тут мета полягає в тому, щоб зменшити число випадків зараження (кількість контактів, середнє число вторинних випадків, які генерує кожен випадок), R - до рівня нижче 1 і, отже, зменшити число випадків до низького рівня або (як у випадку з атиповою пневмонією або лихоманкою Ебола) усунути передачу вірусу від людини до людини.

Основна проблема такого підходу полягає в тому, що нефармацевтичні міри (і ліки, якщо такі є) необхідно підтримувати – в крайньому разі, з перервами - протягом всього періоду циркуляції вірусу в людській популяції або до тих пір, поки не з'явиться вакцина. У випадку з COVID-19 до появи вакцини пройде, принаймні, 12-18 місяців. Крім того, немає гарантії, що початкові вакцини будуть мати високу ефективність.

b) пом'якшення.

В даному випадку мета полягає в тому, щоб використовувати нефармацевтичні міри (а також вакцини або ліки, якщо такі є) не для того, щоб повністю перервати передачу інфекції, а для того, щоб зменшити вплив епідемії на здоров'я, що аналогічно стратегії, прийнятої деякими містами США в 1918 р. і в цілому в світі в 1957, 1968 і 2009 рр. в зв'язку з пандеміями грипу. Під час пандемії 2009 р. наприклад, ранні поставки вакцини були націлені на осіб з раніше існуючими медичними захворюваннями, які піддавали їх ризику виникнення ускладнень та виклику більш важкої форми захворювання. При такому сценарії імунітет населення наростає в результаті епідемії, що в кінцевому підсумку веде до швидкого зниження числа випадків захворювання і зниження рівня передачі інфекції до низьких рівнів.

Ці стратегії відрізняються тим, чи спрямовані вони на скорочення числа випадків зараження R до рівня нижче 1 (пригнічення) і, таким чином, призводять до зниження числа випадків, або просто уповільнюють поширення, знижуючи R , але не до рівня нижче 1.

З самого початку важливо відзначити, що з урахуванням того, що атипова пневмонія (SARS-CoV-2) є новим вірусом, багато що ще належить зрозуміти про його передачу. Крім того, вплив багатьох нефармацевтичних мір, докладно описаних тут, у вирішальній мірі залежить від того, як люди реагують на їх появу, що, швидше за все, буде відрізнятися в різних країнах і навіть спільнотах. І, нарешті, досить імовірно, що відбудуться значні спонтанні зміни в поведінці населення навіть при відсутності санкціонованих урядом заходів втручання. В роботі не розглядаються етичні або економічні

наслідки будь-якої з цих стратегій, за винятком того, що тут немає простого політичного рішення, яке можна було б прийняти.

Пригнічення, яке до цих пір успішно здійснювалося в Китаї і Південній Кореї, пов'язане з величезними соціальними і економічними витратами, які самі по собі можуть викликати істотний вплив на здоров'я і благополуччя в короткостроковій і довгостроковій перспективі. Пом'якшення наслідків ніколи не зможе повністю захистити тих, хто знаходиться в групі ризику, від важких захворювань або смерті, тому пов'язана з цим смертність може залишатися високою. Найкраще сконцентрувати увагу на можливості втілення тоді, приділяючи особливу увагу тому, яким буде можливий вплив цих двох підходів на систему охорони здоров'я.

Нижче отримано результати прогнозування процесу поширення коронавірусу для України, Китаю, Іспанії, Бразилії, США, але вони в рівній мірі застосовні і до більшості країн зі схожим рівнем доходу.

3.2 Побудова математичних моделей процесів поширення коронавірусу

3.2.1 Побудова моделей для прогнозування числа тих, хто захворів в Україні

Для побудови математичних моделей процесів, пов'язаних з поширенням корона-вірусу в Україні скористаємось системою Eviews, яка забезпечує побудову регресійних моделей лінійних та нелінійних нестационарних процесів на основі статистичних даних. Спочатку розглянемо процес накопичення кількості жителів, які захворіли на коронавірус. Дані по захворюваності взято за період з 23.03.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.5:

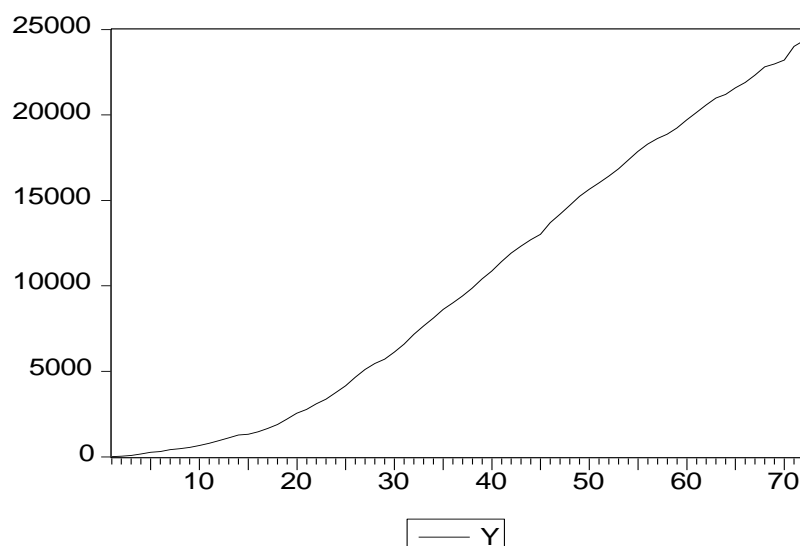


Рисунок 3.5 Захворюваність на корона-вірус в Україні за період з 23.03.2020 по 02.06.2020

На графіку у деяких випадках можна помітити як проявляється присутністю тренду, очевидна періодичність і також якщо не вдається поділити даний часовий ряд на частини зі схожим середнім, що в свою чергу характеризує непостійність математичного сподівання.

Побудувавши корелограму, тобто графічне зображення автокореляційної функції, можна оцінити стаціонарність часового ряду. В разі, коли теоритична АКФ наближається до нуля поступово, тоді ряд вважається стаціонарним для авторегресійного процесу. Коли АКФ зривається і різко стає рівним нулю, ряд являється стаціонарним для процесу ковзного середнього. Хоча й на адекватність визначення стаціонарності впливає наскільки довгий ряд, як повільно збігається АКФ.

Взагалі перевірка стаціонарності впливає з його означення: попередньо розділивши ряд на декілька частин, вводяться гіпотези про рівність математичних сподівань та дисперсій цих частин. У випадку підтвердження сформованих гіпотез при можливості перевіряють також рівність АКФ, тобто використовують спеціальний тест на рівність кореляції попарного порівняння коефіцієнтів кореляції кожного порядку.

Для перевірки стаціонарності застосовують критерії Дікі-Фулера, американського статистика Артура Фулера (1931 року народження) та його учня, теж американського статистика Девіда Алана Дікі (1945 року народження). Ці критерії також відомі, як тест на наявність одиничного кореня:

- 1). Висунути припущення стосовно виду процесу, який утворив даний часовий ряд;
- 2). Побудувати допоміжну модель і перевірити гіпотези щодо коефіцієнтів цієї моделі;
- 3). Зробити висновки про стаціонарність вихідного часового ряду, що досліджувався;

Єдине, що ускладнює використання даного методу та зрештою може вплинути на результат це можлива некоректність складених припущень щодо виду допоміжної моделі і незадатлива потужність тесту для невеликих часових рядів (менше 100 спостережень)

Для визначення порядку авторегресії обчислимо автокореляційну функцію цього процесу, яка подана нижче в табл. 3.1:

Таблиця 3.1 Автокореляційна функція процесу захворюваності

Date: 05/02/20 Time: 20:30
Sample: 1901 1972
Included observations: 72

Autocorrelation	Partial Correlation	A C	PAC	Q- Stat	Prob
. *****	. *****	0.966	0.966	69.998	0.000
. *****	. .	0.931	-0.033	135.91	0.000
. *****	. .	0.896	-0.009	197.89	0.000
. *****	. .	0.860	-0.036	255.86	0.000
. *****	. .	0.823	-0.037	309.74	0.000
. *****	. .	0.786	-0.026	359.57	0.000
. *****	. .	0.748	-0.028	405.40	0.000

. *****	. .	0.	-	447	0.
		709	0.034	.25	000
. *****	. .	0.	-	485	0.
		670	0.031	.18	000

Очевидно, що це нестационарний процес з трендом 1-го порядку, порядок авторегресії дорівнює одиниці. Тому першим рівнянням, яке будемо будувати, буде $AR(1)$. Результати оцінювання рівняння авто регресії першого порядку наведені в табл. 3.2:

Таблиця 3.2 Результати оцінювання рівняння авто регресії першого порядку

Dependent Variable: Y				
Method: Least Squares				
Date: 05/02/20 Time: 20:34				
Sample(adjusted): 1902 1972				
Included observations: 71 after adjusting endpoints				
$Y=C(1)+C(2)*Y(-1)$				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	231.2579	27.28157	8.476709	0.0000
C(2)	1.011200	0.002152	469.9185	0.0000
R-squared	0.999688	Mean dependent var		10285.45
Adjusted R-squared	0.999683	S.D. dependent var		8012.019
S.E. of regression	142.6266	Akaike info criterion		12.78610
Sum squared resid	140362.2	Schwarz criterion		12.84984
Log likelihood	-451.9066	Durbin-Watson stat		0.885576

Рівняння має вигляд:

$$y(k) = 231,258 + 1,0112 y(k-1) \quad (3.6)$$

Основні статистичні характеристики моделі:

$$R^2 = 0,999; \quad DW = 0,856; \quad CKII = 1403622,0.$$

Очевидно, що ця модель є незадовільною за статистикою Дарбіна-Уотсона, яка складає $DW = 0,856$, в той час, коли ідеальне значення цієї статистики складає 2,0. Знайдемо характеристики одно крокового прогнозу, які зображені нижче на рисунку 3.6:

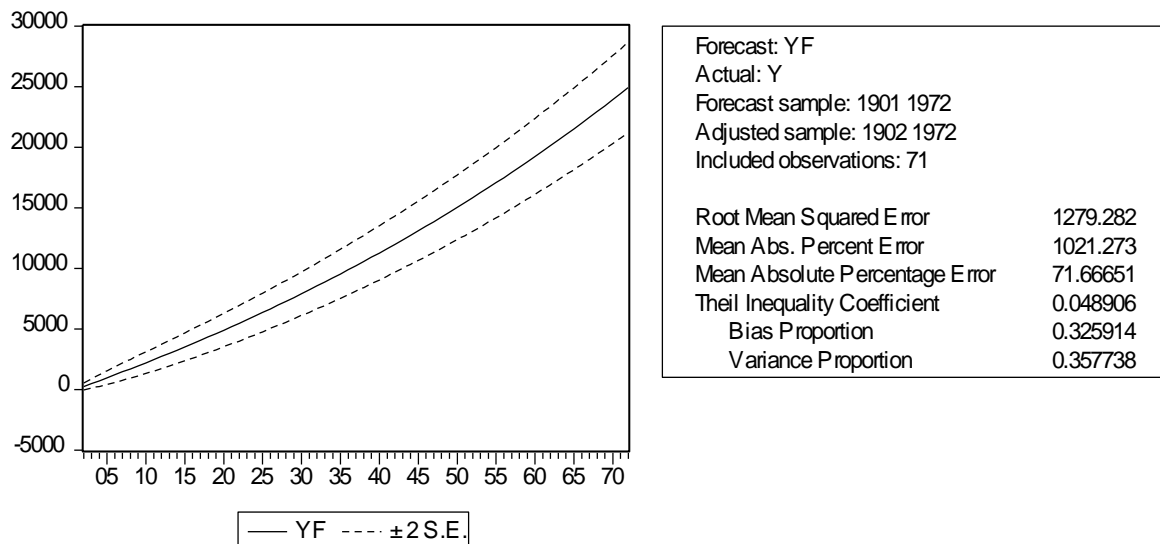


Рисунок 3.6 Характеристики однокрокового прогнозу за моделлю $AR(1)$

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 1279,282; \quad CAП = 1021,273; \quad CAПП = 71,67; \quad U = 0,0489.$$

За допомогою даної моделі неможливо отримати прогноз високої якості, тому що середня абсолютна похибка у процентах (САПП або MAPE) склала $CAПП = 71,67\%$.

Оскільки процес містить тренд щонайменше 1-го порядку, то оцінимо модель з трендом. Така модель для квадратичного тренду має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + c_1 k + c_2 k^2 + \varepsilon(k), \quad (3.7)$$

де $k = 0, 1, 2, \dots$ – дискретний час.

Результати оцінювання моделі з квадратичним трендом подані у таблиці 3.3:

Таблиця 3.3 Результати оцінювання моделі $AP(1) +$ тренд другого порядку

Dependent Variable: Y				
Method: Least Squares				
Date: 05/02/20 Time: 20:54				
Sample(adjusted): 1902 1972				
Included observations: 71 after adjusting endpoints				
$Y = C(1) + C(2) * Y(-1) + C(3) * K + C(4) * K^2$				
	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-94.32666	48.90131	-1.928919	0.0580
C(2)	0.981964	0.016732	58.68736	0.0000
C(3)	25.88876	4.107055	6.303485	0.0000
C(4)	-0.190947	0.051693	-3.693900	0.0004
R-squared	0.999851	Mean dependent var	10285.45	
Adjusted R-squared	0.999844	S.D. dependent var	8012.019	
S.E. of regression	100.0424	Akaike info criterion	12.10375	
Sum squared resid	670568.4	Schwarz criterion	12.23123	
Log likelihood	-425.6833	Durbin-Watson stat	1.792133	

Рівняння має вигляд:

$$y(k) = -94,327 + 0,982 y(k-1) + 25,89 k - 0.191 k^2 \quad (3.8)$$

Основні статистичні характеристики моделі такі:

$$R^2 = 0,999; DW = 1,792; CKII = 670568,4.$$

Статистика Дарбіна-Уотсона значно покращилась і склала: $DW = 1,792$, тобто вона суттєво наблизилась до ідеального значення. На рис. 3.7 зображено однокроковий прогноз за моделлю $AP(1) +$ тренд другого порядку:

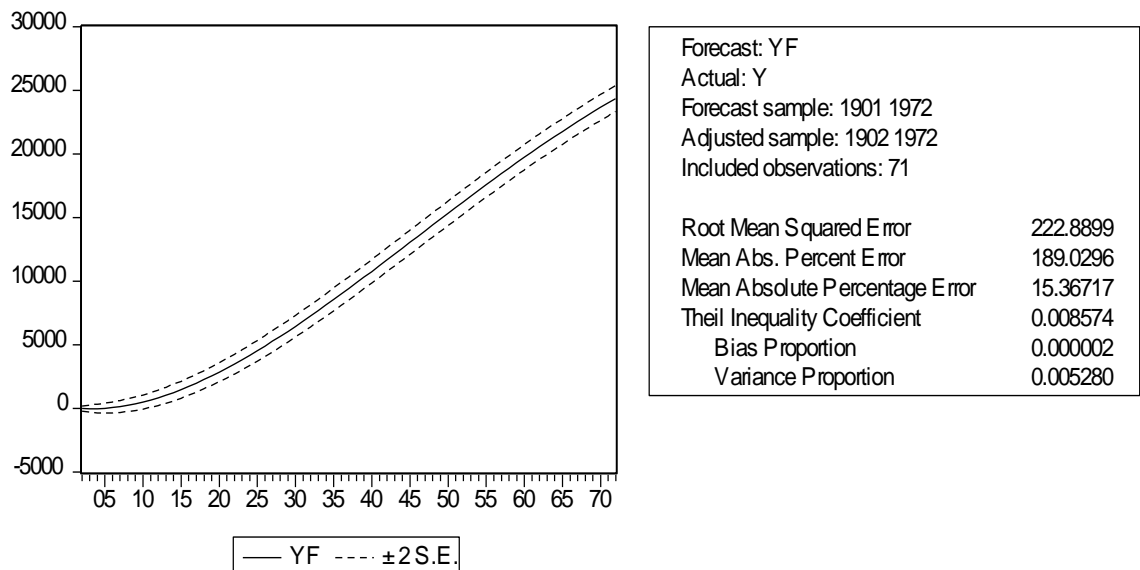


Рисунок 3.7 Характеристики однокрокового прогнозу за моделлю $AR(1)$ + тренд другого порядку

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 222,89; \quad CAП = 189,03; \quad CAПП = 15,37; \quad U = 0,0086.$$

Тобто характеристики однокрокового прогнозу значно покращились. За допомогою даної моделі можна отримати прогноз високої якості, тому що середня абсолютна похибка у процентах (САПП або MAPE) склала: САПП = 15,37 %. Зведено побудовані моделі в одну таблицю (табл. 3.4), подану нижче.

Таблиця 3.4 Зведена таблиця для двох побудованих моделей захворюваності

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
AR(1)	0,999	1403622	0,856	1279,	1021,	71,67	0,0489
AR(1)+t2	0,999	670568	1,792	223,0	189,3	15,37	0,0086

Таблиця 3.4 наочно свідчить про суттєве покращення адекватності моделі з трендом і про підвищення якості одно крокового прогнозу за цією ж моделлю з трендом. Подальшим покращенням прогнозу може бути застосування моделі авторегресії з ковзним середнім, але це потрібно перевіряти.

Побудуємо АКФ та ЧАКФ для остатків моделі $AR(1)$, щоб визначити порядок ковзного середнього для побудови моделі $ARCS(1, q)$, результати представлені в табл. 3.5:

Таблиця 3.5 АКФ остатків моделі $AR(1)$

Date: 05/05/20 Time: 22:14
Sample (adjusted): 2 93
Included observations: 92 after adjustments

Autocorrelation	Partial Correlation	A C	P AC	Q-Stat	Pr ob
		0.	0.	15.	0.
. ***	. ***	401	401	275	000
		0.	0.	86.	0.
. *****	. *****	858	831	009	000
		0.	-	100	0.
. ***	. .	389	0.009	.71	000
		0.	0.	156	0.
. *****	. .	753	055	.45	000
		0.	0.	175	0.
. ***	. **	436	241	.35	000
		0.	0.	227	0.
. *****	. *	722	211	.80	000
		0.	0.	250	0.
. ***	. .	475	070	.72	000
		0.	-	297	0.
. *****	. .	671	0.003	.10	000
		0.	-	313	0.
. ***	** .	402	0.306	.92	000
		0.	-	347	0.
. *****	** .	0	568 0.228	.93	000
		0.	0.	363	0.
. ***	. *	1	384 096	.70	000
		0.	-	389	0.
. *****	.* .	2	489 0.168	.56	000
		0.	-	407	0.
. ***	. .	3	407 0.008	.72	000
		0.	0.	430	0.
. ***	. *	4	454 129	.53	000
		0.	-	447	0.
. ***	. .	5	386 0.040	.23	000

			0.	-	461	0.
.	***		.	***		6
			360	0.107	.98	000
			0.	-	471	0.
.	**		.	**		7
			294	0.119	.96	000
			0.	-	481	0.
.	**		.	**		8
			287	0.096	.61	000
			0.	-	489	0.
.	**		.	*		9
			252	0.003	.14	000

Кореляція присутня на 1, 2, 5, 9, 10 лазі при граничному значенні 0,2.

Результати оцінювання параметрів моделі АРКС(1, 1) представлені в таблиці 3.6 нижче:

Таблиця 3.6 Результати оцінювання моделі авторегресії з ковзним середнім першого порядку

Dependent Variable: Z				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 05/05/20 Time: 22:13				
Sample: 1 93				
Included observations: 93				
Convergence achieved after 15 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
	12164.36	12090.49	1.006109	0.3171
C	0.999577	0.005807	172.1440	0.0000
AR(1)	0.847786	0.055641	15.23669	0.0000
MA(1)	36586.69	6740.591	5.427816	0.0000
SIGMASQ				
R-squared	0.999459			7864.989
Adjusted R-squared	0.999441			8269.650
S.E. of regression	195.5276			13.53425
Sum squared resid	340256.2			13.64318
Log likelihood	625.3426			823.0525
F-statistic	41.00000			738
Prob(F-statistic)	00			

Рівняння має вигляд:

$$y(k) = 12164,36 + 0,999y(k-1) + mv(k) + 0,848mv(k-1) \quad (3.9)$$

Основні статистичні характеристики моделі такі:

$$R^2 = 0,999; DW = 0,526; СКП = 3402562.$$

Очевидно, що ця модель є незадовільною за статистикою Дарбіна-Уотсона, яка складає $DW = 0,526$, в той час, коли ідеальне значення цієї статистики складає 2,0. Розглянемо показники якості побудованого прогнозу моделі $АРКС(1, 1)$, які зображені нижче на рисунку 3.8:

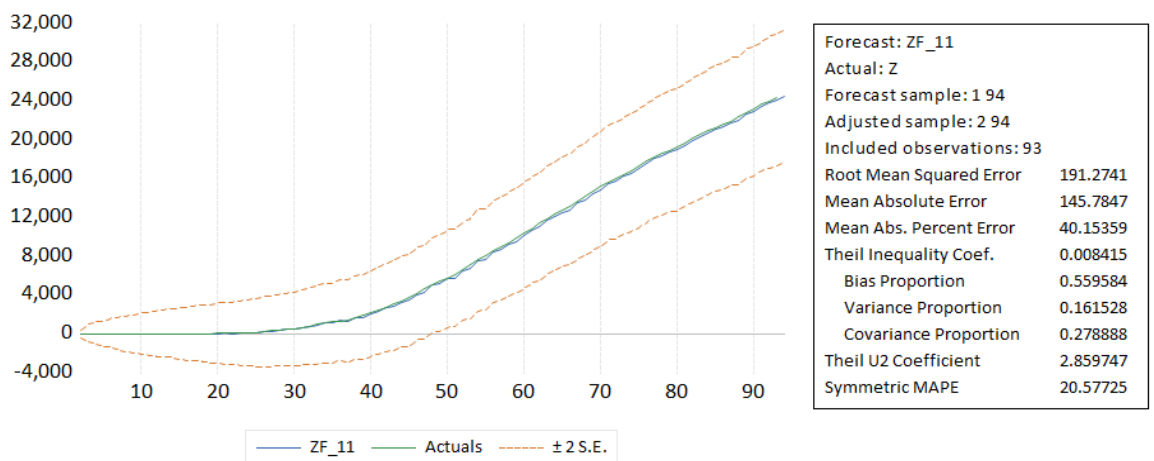


Рисунок 3.8 Характеристики однокрокового прогнозу за моделлю $АРКС(1,1)$

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 191,274; САП = 145,785; САПП = 40,15; U = 0,0084.$$

За допомогою даної моделі неможливо отримати прогноз високої якості, тому що середня абсолютна похибка у відсотках (САПП або MAPE) склала: $САПП = 40,15 \%$. Це гірше результат, ніж ми отримали при $АР(1)$ з квадратичним трендом.

Тому продовжуємо будувати моделі АРКС(1, q) та їх прогнози для q=2, 10. Результати побудованих моделей зведено в одну таблицю (табл. 3.7), подану нижче.

Таблиця 3.7 Зведена таблиця для побудованих моделей АРКС(1,q) випадків захворювання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АРКС(1, 1)	0,999	3402562	0,526	191,274	145,78	40,15	0,0084
АРКС(1, 2)	0,999	1331058	0,7004	121,6	90,47	30,77	0,0053
АРКС(1, 3)	0,999	933986	1,7562	101,9	72,74	34,88	0,0045
АРКС(1, 4)	0,999	708350	1,4384	88,73	61,86	29,16	0,0039
АРКС(1, 5)	0,999	520113	1,8633	76,04	54,87	32,15	0,0033
АРКС(1, 6)	0,999	502081	1,8915	74,71	50,17	32,83	0,0033
АРКС(1, 9)	0,999	316121	1,8423	61,09	43,80	28,58	0,0027
АРКС(1, 10)	0,999	296938	1,8894	59,23	42,11	29,99	0,0026

Найкращими за показниками якості, а саме значення САПП = 28,58%, виявилась модель АРКС(1, 9), при чому з досить високим значенням статистики Дарвіна-Уотсона DW=1,8423. Параметри рівняння АРКС(1, 9):

$$\begin{aligned}
 y(k) = & 12329,27 + 0,998y(k-1) + mv(k) + 1,1277mv(k-1) \\
 & + 1,3501mv(k-2) + 1,1226mv(k-3) + 1,1911mv(k-4) \\
 & + 1,1949mv(k-5) + 1,1786mv(k-6) + 1,3712mv(k-7) \\
 & + 1,0131mv(k-8) + 0,2683mv(k-9)
 \end{aligned} \quad (3.10)$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 9), які зображені нижче на рисунку 3.9:

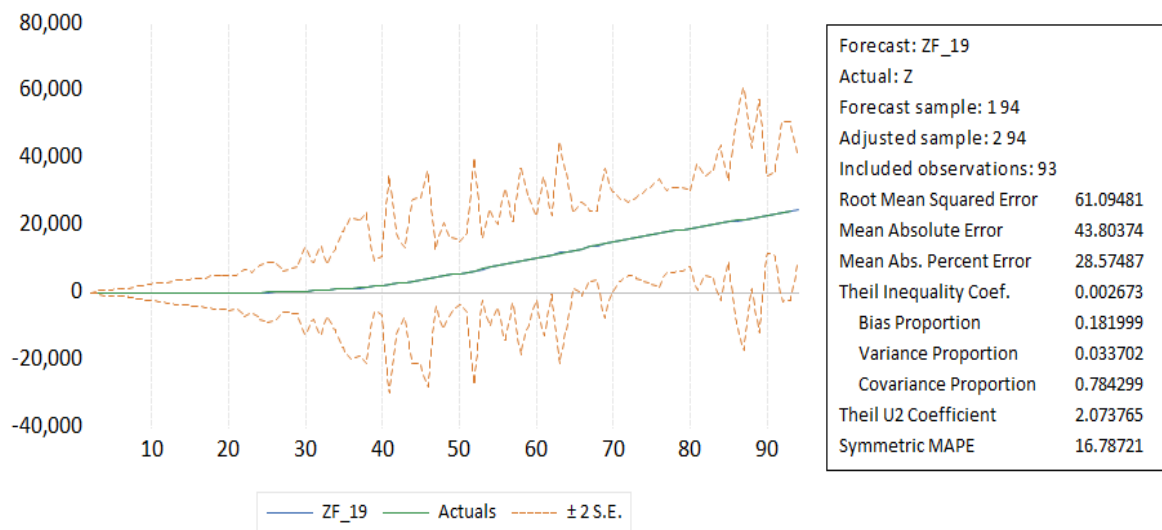


Рисунок 3.9 Характеристики однокрокового прогнозу за моделлю АРКС(1,9)

Отож, очікуваного покращення якості прогнозу не відбулося при застосуванні моделі авто регресійного рівняння з ковзним середнім.

Для порівняння розглянемо можливість моделювання складного нелінійного процесу Торнквіста, який апроксимуємо вибраними лінійними моделями. Найкращий одно кроковий прогноз характеризується значенням САПП = 34,8%, тобто якість прогнозу недостатня. Результати цього дослідження подані у табл. 3.8:

Таблиця 3.8 Результати апроксимації нелінійного процесу Торнквіста лінійними моделями

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АР6	0,34	9542,2	1,993	11,84	9,17	141,97	0,4
АР13	0,34	8107,2	1,95	11,34	9,01	152,22	0,38
АР6 КС1	0,94	901,97	1,55	4,38	3,49	54,4	0,13

AP8 KC3	0,98	274,24	2,84	2,97	2,53	34,8	0,088
---------	------	--------	------	------	------	------	-------

Порівняння таблиць 3.7 і 3.8 свідчить про те, що модель, побудована для захворюваності має набагато кращі характеристики ніж модель нелінійного процесу Торнквіста. Тобто існує принципіальна можливість покращити результати моделювання і прогнозування.

3.2.2 Побудова моделей для прогнозування числа тих, хто одужав в Україні

Тепер розглянемо процес, пов'язаний з кількістю тих, хто вилікувався в Україні протягом того ж самого періоду часу, а саме: з 23.03.2020 по 02.06.2020. Для цього скористаємось ще однією вибіркою даних, яка характеризує цей процес. Кількість значень склала 68. Графік процесу подано на рис. 3.10.

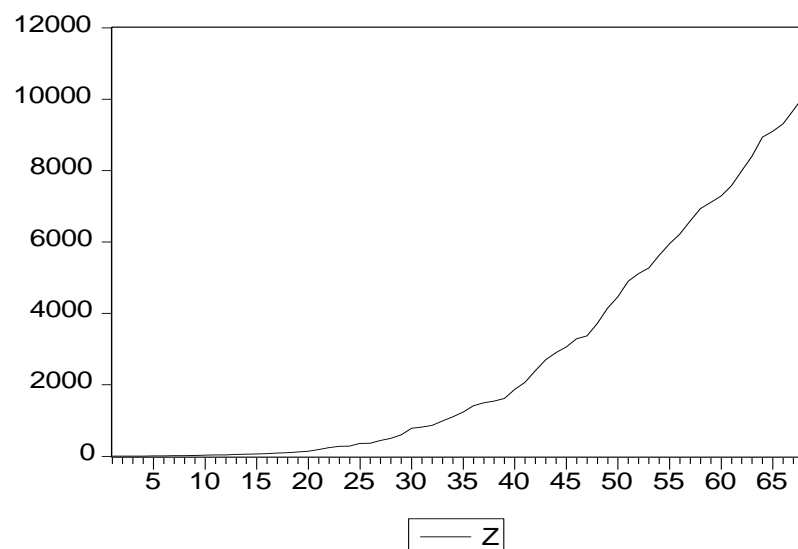


Рисунок 3.10 Графік процесу видужування після зараження вірусом

Автокореляційна функція цього процесу подана у таблиці 3.9:

Таблиця 3.9 АКФ процесу видужування

Date: 05/03/20 Time: 18:56

Sample: 1 68

Included observations: 68

Autocorrelation	Partial Correlation	A C	PAC	Q- Stat	Prob	
. *****	. *****	0.	0.	64.	0.	
		951	951	196	000	
. *****	. .	0.	-	122	0.	
		902	0.021	.82	000	
. *****	. .	0.	-	176	0.	
		853	0.019	.15	000	
. *****	. .	0.	-	224	0.	
		804	0.039	.20	000	
. *****	. .	0.	-	267	0.	
		753	0.044	.02	000	
. *****	. .	0.	-	305	0.	
		704	0.003	.09	000	
. *****	. .	0.	-	338	0.	
		657	0.017	.76	000	
. *****	. .	0.	-	368	0.	
		611	0.013	.37	000	
. *****	. .	0.	-	394	0.	
		565	0.027	.14	000	
. *****	. .	0.	-	416	0.	
		0	518	0.041	.20	000
. *****	. .	0.	-	434	0.	
		1	471	0.040	.70	000

Очевидно, що це нестационарний процес з трендом 1-го порядку, порядок авторегресії дорівнює одиниці. Тому першим рівнянням, яке будемо будувати AR(1), результати оцінювання представлені в табл. 3.10:

Таблиця 3.10 Результати оцінювання моделі авторегресії першого порядку з трендом

Dependent Variable: Z

Method: Least Squares

Date: 06/03/20 Time: 19:03

Sample(adjusted): 2 68

Included observations: 67 after adjusting endpoints

 $Z = C(1) + C(2) * Z(-1) + C(3) * K + C(4) * K^2$

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	46.1698	44.07454	1.047541	0.2989

C(2)	0.86581	0.049361	17.54049	0.0000
	3			
C(3)	-	5.721663	-	0.0631
	10.82398		1.891754	
C(4)	0.51760	0.178715	2.896266	0.0052
	7			
R-squared	0.99946	Mean dependent var	2718.3	
	1		28	
Adjusted R-squared	0.99943	S.D. dependent var	3135.6	
	5		99	
S.E. of regression	74.5041	Akaike info criterion	11.517	
	4		43	
Sum squared resid	349704.	Schwarz criterion	11.649	
	6		05	
Log likelihood	-	Durbin-Watson stat	1.4666	
	381.8340		74	

Рівняння має такий вигляд:

$$z(k) = 46,17 + 0,866z(k-1) - 10,82k + 0.517k^2 \quad (3.11)$$

Основні статистичні характеристики моделі:

$$R^2 = 0,999; DW = 1,47; CKП = 349704,6.$$

Очевидно, що ця модель є недостатньо задовільною за статистикою Дарбіна-Уотсона, яка складає $DW = 1,47$, в той час, коли ідеальне значення цієї статистики складає 2,0. Знайдемо характеристики однокрокового прогнозу, які зображені нижче на рисунку 3.11:

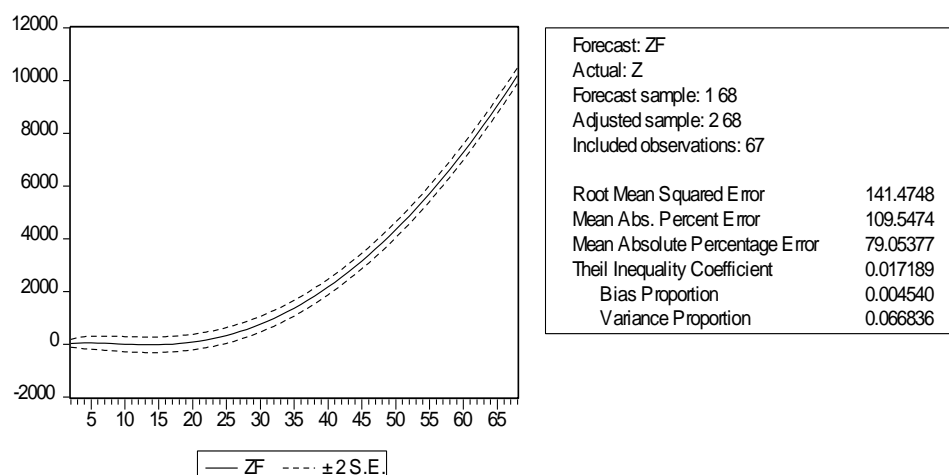


Рисунок 3.11 Характеристики однокрокового прогнозу

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 141,47; \quad CAI = 109,55; \quad CAII = 79,05; \quad U = 0,017.$$

Оскільки $CAII = 79,05$, то характеристики одно крокового прогнозу незадовільні. Потрібно спробувати побудувати кращу модель.

Наприклад, спробуємо побудувати модель $APKC(1, q)$, використовуючи варіант побудови ковзного середнього по залишкам $AP(1)$ рівняння моделі. Для визначення порядку ковзного середнього побудуємо АКФ та ЧАКФ для остатків моделі $AP(1)$, табл. 3.10. Легко помітити, що при граничному значенні 0,3 остатки між собою не корелюють. Результати наведено в табл.3.11:

Таблиця 3.11 Побудова АКФ та ЧАКФ для остатків моделі $AP(1)$

Date: 05/05/20 Time: 16:30
Sample: 1 68
Included observations: 68

Autocorrelation	Partial Correlation	A C	P AC	Q-Stat	Prob
		0.	0.	39.	0.
. *****	. *****	745	745	407	000
		0.	0.	67.	0.
. *****	. *.	629	167	946	000
		0.	0.	96.	0.
. ****	. **	622	247	238	000
		0.	0.	121	0.
. ****	. *.	584	076	.57	000
		0.	0.	148	0.
. ****	. *.	600	200	.78	000
		0.	0.	182	0.
. *****	. **	662	258	.43	000
		0.	-	209	0.
. ****	. .	592	0.060	.79	000
		0.	-	232	0.
. ****	. .	539	0.005	.80	000
		0.	-	253	0.
. ****	. .	502	0.062	.16	000
		0.	-	269	0.
. ***	. * .	0	442	0.092	.21 000
		0.	-	281	0.
. ***	. * .	1	390	0.138	.89 000
		0.	0.	296	0.
. ***	. .	2	418	035	.73 000
		0.	0.	314	0.
. ***	. *.	3	458	153	.86 000

			0.	-	327	0.
. ***		. .	4	385	0.118	.93 000
				0.	-	337 0.
. **		. .	5	333	0.039	.88 000
				0.	-	344 0.
. **		. .	6	265	0.120	.29 000
				0.	-	347 0.
. *		. .	7	182	0.118	.37 000
				0.	0.	351 0.
. *		. .	8	204	0.033	.33 000
				0.	-	354 0.
. *		. .	9	177	0.161	.38 000
				0.	-	355 0.
. *		. .	0	127	0.026	.99 000

Але якщо зменшити граничне значення до 0,2, то кореляція буде присутня на 1, 3, 6 лазі. Тому першим рівнянням, яке будемо будувати, буде АРКС(1, 1), результати наведено в табл.3.12:

Таблиця 3.12 Результати оцінювання рівняння авто регресії із ковзним середнім першого порядку

Dependent Variable: Z				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 06/05/20 Time: 16:49				
Sample: 1 68				
Included observations: 68				
Convergence achieved after 13 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5045.226	5026.226	1.003780	0.3193
AR(1)	0.998826	0.011193	89.23956	0.0000
MA(1)	0.748249	0.070470	10.61800	0.0000
SIGMASQ	19658.79	3185.097	6.172117	0.0000
R-squared	0.997963	Mean dependent var		2678.368
Adjusted R-squared	0.997867	S.D. dependent var		3129.607
S.E. of regression	144.5250	Akaike criterion		12.95932
Sum squared resid	133679.8	Schwarz criterion		13.08988
Log likelihood	436.6170	Hannan-Quinn criter.		13.01106

	10451.		0.999
F-statistic	10	Durbin-Watson stat	984
	0.0000		
Prob(F-statistic)	00		

Рівняння має вигляд:

$$y(k) = 5045,23 + 0,999y(k-1) + mv(k) + 0,748mv(k-1) \quad (3.12)$$

Основні статистичні характеристики моделі:

$$R^2 = 0,999; DW = 0,9999; СКП = 1336798,0.$$

Очевидно, що ця модель є незадовільною за статистикою Дарбіна-Уотсона, яка складає $DW = 0,9999$. Знайдемо характеристики одно крокового прогнозу, які зображені нижче на рисунку 3.12:

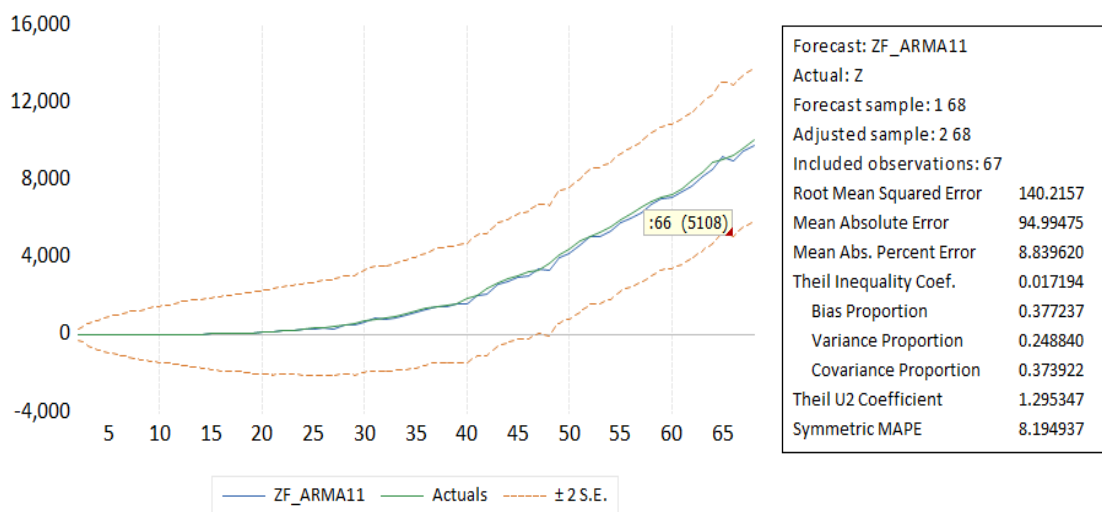


Рисунок 3.12 Характеристики однокрокового прогнозу за моделлю $APRC(1,1)$

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 140,216; САП = 94,994; САПП = 8,839; U = 0,017.$$

За допомогою даної моделі можна отримати прогноз досить високої якості, тому що середня абсолютна похибка у процентах (САПП або MAPE)

склала: САПП = 8,83 %. Ця модель суттєво відрізняється показниками якості прогнозу від АР(1) з трендом.

Для рівняння АРКС(1, 2) результати представлені в табл. 3.13:

Таблиця 3.13 Результати оцінювання рівняння авто регресії із ковзним середнім другого порядку

Dependent Variable: Z				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 05/05/20 Time: 17:37				
Sample: 1 68				
Included observations: 68				
Convergence achieved after 16 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5010.19	4986.741	1.004688	0.3189
AR(1)	0.998315	0.013440	74.28212	0.0000
MA(1)	0.987843	0.104305	9.470755	0.0000
MA(2)	0.389623	0.123547	3.153652	0.0025
SIGMASQ	15080.64	2182.642	6.909351	0.0000
R-squared	0.998437	Mean dependent var		2678.368
Adjusted R-squared	0.998338	S.D. dependent var		3129.607
S.E. of regression	127.5834	Akaike info criterion		12.73046
Sum squared resid	102548.3	Schwarz criterion		12.89366
Log likelihood	-427.8358	Hannan-Quinn criter.		12.79513
F-statistic	10062.99	Durbin-Watson stat		1.636490
Prob(F-statistic)	0.000000			

Рівняння має вигляд:

$$y(k) = 5010,12 + 0,998y(k - 1) + mv(k) + 0,987mv(k - 1) + 0,389mv(k - 2) \quad (3.13)$$

Основні статистичні характеристики моделі:

$$R^2 = 0,998; DW = 1,636; СКП = 1025483.$$

Знайдемо характеристики однокрокового прогнозу, які зображені нижче на рисунку 3.13

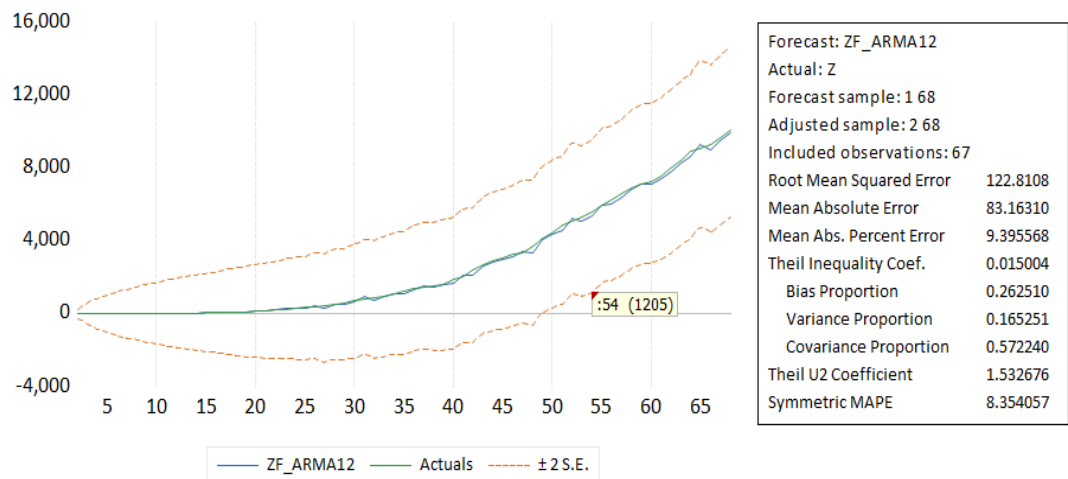


Рисунок 3.13 Характеристики однокрокового прогнозу за моделлю $APRC(1,2)$

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 122,81; САП = 83,16; САПП = 9,395; U = 0,015.$$

Тож ми спостерігаємо покращення значення статистики Дарбіна-Уотсона, в результаті чого побудований нам прогноз високої якості побудувати не вдасться.

Зведемо побудовані моделі в одну таблицю (табл. 3.14), подану нижче.

Таблиця 3.14 Зведена таблиця для двох побудованих моделей ситуації з тими, хто одужав

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла

АРКС(1, 1)	0,999	133679	0,999	140,216	94,994	8,839	0,017
АРКС(1, 2)	0,998	102548	1,636	122,81	83,16	9,395	0,015
АРКС(1, 3)	0,998	691256	1,693	100,9	65,96	9,251	0,012
АРКС(1, 4)	0,999	672017	1,958	99,42	64,63	10,11	0,012
АРКС(1, 5)	0,999	649515	1,882	97,93	68,86	11,702	0,011
АРКС(1, 6)	0,999	514097	1,759	88,69	58,32	9,196	0,011

Таблиця 3.13 наочно свідчить про суттєве підвищення якості однокрокового прогнозу за моделлю авто регресійного рівняння з ковзним середнім. Найкращою серед усіх розглянутих з показником середня абсолютна похибка у відсотках (САПП або MAPE) САПП = 8,83 % модель АРКС(1, 1).

3.2.3 Побудова моделей для прогнозування числа тих, хто захворів в Китаї

Китай обраний не випадково, адже це країна, з якої все почалося. Саме тут були зареєстровані перші випадки захворювання і на момент створення статистики їхня кількість уже складала більше п'яти сотень чоловік. Проте, завдяки ефективним та швидким методам локалізації епідемії, дисциплінованості населення і не без диктаторської політики комуністичної партії, яка керує країною, Китай на даний момент майже зупинив поширення коронавірусу на своїй території, не перевищивши навіть 90 000 зареєстрованих випадків інфікування.

Спочатку розглянемо процес накопичення кількості жителів, які захворіли на коронавірус. Дані по захворюваності взято за період з 22.01.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.14.

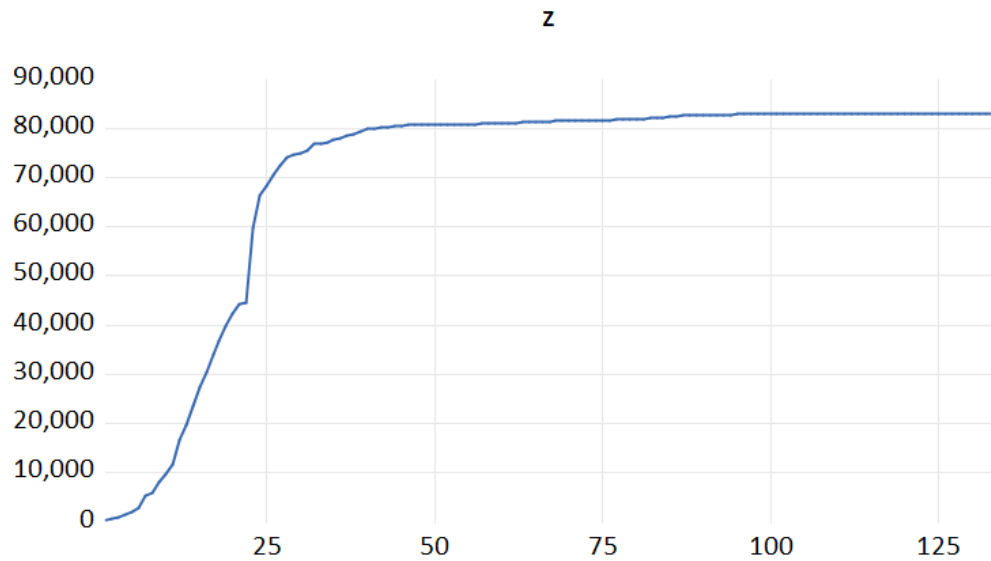


Рисунок 3.14 Захворюваність на коронавірус в Китаї за період з
22.01.2020 по 02.06.2020

Для визначення порядку авторегресії обчислимо автокореляційну функцію цього процесу, яка подана нижче табл. 3.15.

Таблиця 3.15 Автокореляційна функція процесу захворюваності

Date: 05/07/20 Time: 20:16
Sample: 1 133
Included observations: 133

Autocorrelation	Partial Correlation	A C	P AC	Q-Stat	Prob
1.000000	1.000000	0.	0.	126	0.
0.965000	0.965000	965	965	.64	000
0.244000	0.244000	0.	-	244	0.
0.927000	0.927000	927	0.063	.34	000
0.352000	0.352000	0.	-	352	0.
0.886000	0.886000	886	0.053	.78	000
0.451000	0.451000	0.	-	451	0.
0.843000	0.843000	843	0.052	.73	000
0.541000	0.541000	0.	-	541	0.
0.798000	0.798000	798	0.051	.09	000
0.620000	0.620000	0.	-	620	0.
0.751000	0.751000	751	0.048	.89	000
0.691000	0.691000	0.	-	691	0.
0.704000	0.704000	704	0.022	.61	000
0.753000	0.753000	0.	-	753	0.
0.656000	0.656000	656	0.050	.41	000
0.806000	0.806000	0.	-	806	0.
0.607000	0.607000	607	0.032	.75	000
0.852000	0.852000	0.	-	852	0.
0.557000	0.557000	0	557	0.041	.05 000

			0.	-	889	0.
. ****		. .	1	507	0.029	.92 000
				0.	0.	921 0.
. ***		. .	2	460	000	.27 000
				0.	-	946 0.
. ***		. .	3	413	0.017	.79 000
				0.	-	967 0.
. ***		. .	4	369	0.002	.30 000
				0.	-	983 0.
. **		. .	5	326	0.006	.49 000
				0.	-	996 0.
. **		. .	6	286	0.010	.03 000
				0.	-	100 0.
. **		. .	7	247	0.009	5.5 000
				0.	-	101 0.
. *		. .	8	210	0.015	2.4 000
				0.	-	101 0.
. *		. .	9	175	0.009	7.2 000
				0.	-	102 0.
. *		. .	0	142	0.013	0.4 000

Очевидно, що це нестационарний процес з трендом 1-го порядку, порядок авторегресії дорівнює одиниці. Тому першим рівнянням, яке будемо будувати, буде $AR(1)$, результати представлено в табл. 3.16:

Таблиця 3.16 Результати оцінювання рівняння авто регресії першого порядку

Dependent Variable: Z				
Method: ARMA Maximum Likelihood (BFGS)				
Date: 05/07/20 Time: 20:18				
Sample: 1 133				
Included observations: 133				
Convergence achieved after 14 iterations				
Coefficient covariance computed using outer product of gradients				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	43515.97	45345.18	0.959660	0.3390
AR(1)	0.999038	0.010144	98.48596	0.0000
SIGMASQ	3191747.	89619.47	35.61443	0.0000
R-squared	0.994559	Mean dependent var		70547.61
Adjusted R-squared	0.994476	S.D. dependent var		24312.39
S.E. of regression	1807.042	Akaike info criterion		17.90609

	4.25E+		17.97
Sum squared resid	08	Schwarz criterion	129
	-	Hannan-Quinn	17.93
Log likelihood	1187.755 criter.		258
	11882.		0.816
F-statistic	10	Durbin-Watson stat	767
	0.0000		
Prob(F-statistic)	00		

Рівняння має вигляд:

$$y(k) = 43515,97 + 0,999y(k - 1) \quad (3.14)$$

Основні статистичні характеристики моделі:

$$R^2 = 0,995; DW = 0,817; СКП = 4,25 \cdot 10^8.$$

Очевидно, що ця модель є незадовільною за статистикою Дарбіна-Уотсона, яка складає $DW = 0,817$, в той час, коли ідеальне значення цієї статистики складає 2,0. Знайдемо характеристики одно крокового прогнозу, які зображені нижче на рисунку 3.15:

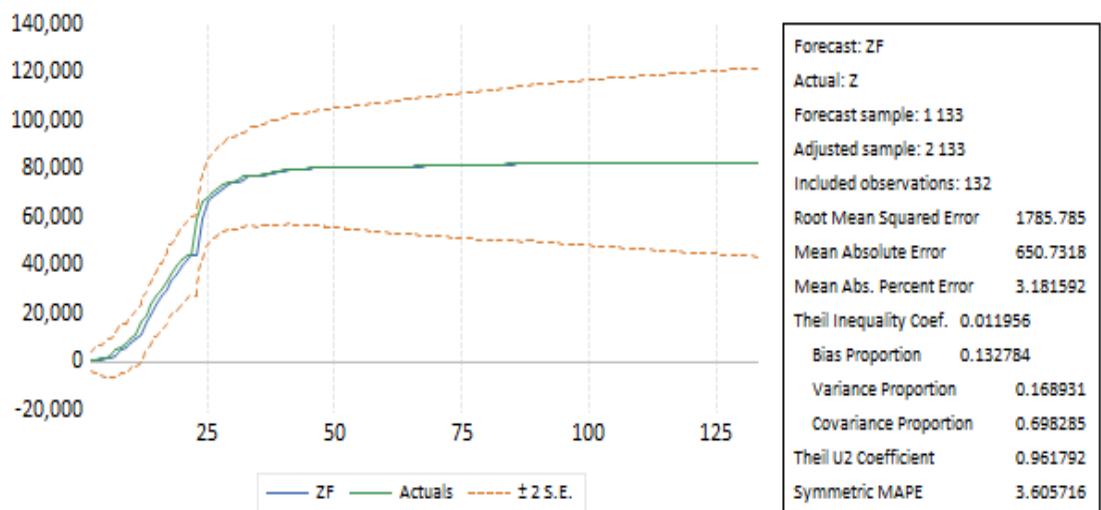


Рисунок 3.15 Характеристики однокрокового прогнозу за моделлю $AR(1)$

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 1785,785; \text{САП} = 650,732; \text{САПП} = 3,1816; U = 0,01196.$$

За допомогою даної моделі отримано прогноз високої якості, тому що середня абсолютна похибка у відсотках (САПП або MAPE) склала: САПП = 71,67 %. Проте, його навряд чи назвеш доволі точним. Оскільки процес містить тренд щонайменше 1-го порядку, то оцінимо модель з трендом. Така модель для квадратичного тренду має вигляд:

$$y(k) = a_0 + a_1 y(k-1) + c_1 k + c_2 k^2 + \varepsilon(k) \quad (3.15)$$

де $k = 0, 1, 2, \dots$ – дискретний час.

Результати оцінювання моделі з квадратичним трендом подані у таблиці 3.17:

Таблиця 3.17 Результати оцінювання моделі $AR(1)$ + тренд другого порядку

Dependent Variable: Z				
Method: Least Squares (Gauss-Newton / Marquardt steps)				
Date: 05/07/20 Time: 20:33				
Sample (adjusted): 2 133				
Included observations: 132 after adjustments				
$Z = C(1) + C(2) * Z(-1) + C(3) * K + C(4) * K^2$				
	Coefficient	Std. Error	t-Statistic	Prob.
	2999.5	405.517	7.39685	0.000
C(1)	59	9	9	0
	0.9933	0.01290	77.0035	0.000
C(2)	69	0	2	0
	-	28.2114	-	0.078
C(3)	50.10579	1	1.776083	1
	0.2454	0.17071	1.43760	0.153
C(4)	26	9	0	0
<hr/>				
	0.9962			71077
R-squared	98	Mean dependent var		.92
Adjusted R-squared	0.9962			23620
	11	S.D. dependent var		.24
	1453.9	Akaike info criterion		17.43
S.E. of regression	58			180
	2.71E+			17.51
Sum squared resid	08	Schwarz criterion		916

	-	Hannan-Quinn	17.46
Log likelihood	1146.499 criter.		730
	11481.		1.261
F-statistic	65	Durbin-Watson stat	111
	0.0000		
Prob(F-statistic)	00		

Рівняння має вигляд:

$$y(k) = 2999,56 + 0,993y(k - 1) - 50,11k - 0,245k^2 \quad (3.16)$$

Основні статистичні характеристики моделі такі:

$$R^2 = 0,996; DW = 1,261; СКП = 2,71 \cdot 10^8.$$

Статистика Дарбіна-Уотсона не набагато, але покращилась і склала: $DW = 1,261$, тобто зрештою вона наблизилась до свого ідеального значення, характеристики побудованого прогнозу представлено на рис.3.16:

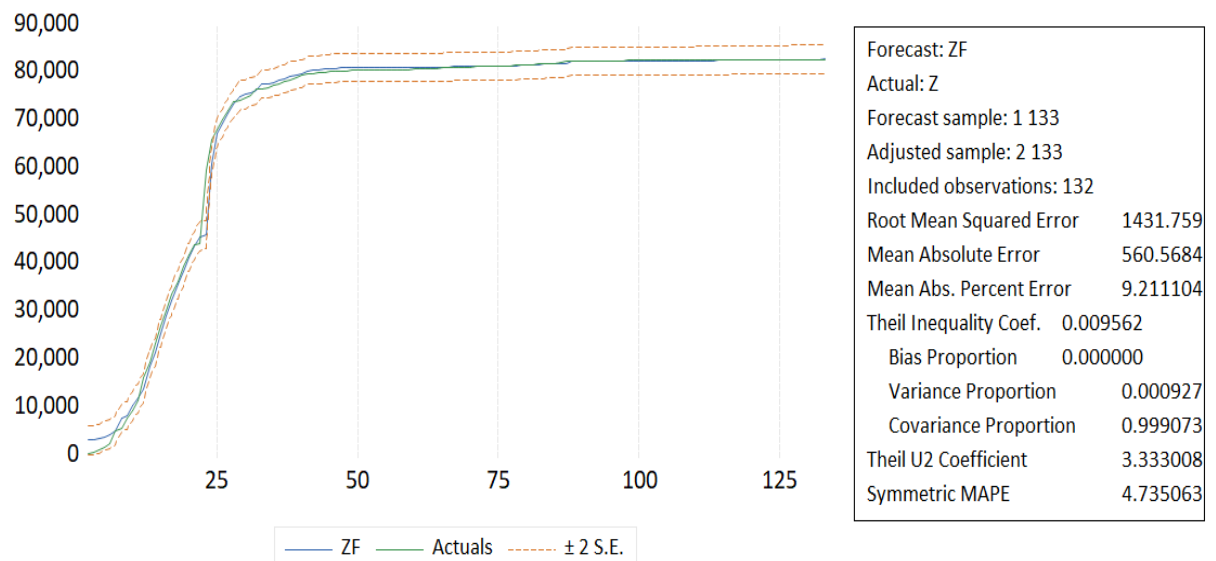


Рисунок 3.16 Характеристики однокрокового прогнозу за моделлю $AR(1)$
+ тренд другого порядку

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 1431,76; САП = 560,568; САПП = 9,21; U = 0,0096.$$

Зважаючи на характеристики прогнозу, що отримано за допомогою моделі авторегресії першого порядку з квадратичним трендом, точність моделі вдалося підвищити, хоча і відносно першого випадку погіршилась якість прогнозу для цієї моделі. За допомогою даної моделі можна отримати прогноз високої якості, тому що середня абсолютна похибка у процентах (САПП або МАРЕ) склала: САПП = 9,21 %. Зведено побудовані моделі в одну таблицю (табл. 3.18), подану нижче.

Таблиця 3.18 Зведена таблиця для двох побудованих моделей захворюваності

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
AP(1)	0,995	4,25*10 ⁸	0,817	1785,78	650,73	3,181	0,01196
AP(1)+t2	0,996	2,71*10 ⁸	1,261	1431,76	560,56	9,21	0,0096

Подальшим покращенням прогнозу може бути застосування моделі авто регресії з ковзним середнім, але це потрібно перевіряти.

Побудуємо АКФ та ЧАКФ для остатків моделі AP(1), щоб визначити порядок ковзного середнього для побудови моделі АРКС(1, q), результати представлені в табл. 3.19:

Таблиця 3.19 АКФ остатків моделі AP(1)

Date: 05/07/20 Time: 21:17						
Sample (adjusted): 2 133						
Included observations: 132 after adjustments						
Autocorrelation	Partial	A	P	Q-	Pr	
n	Correlation	C	AC	Stat	ob	
		0.	0.	17.	0.	
.***	.***	355	355	005	000	
		0.	0.	21.	0.	
. *	. .	181	063	453	000	
		0.	0.	26.	0.	
. *	. *	197	132	784	000	
		0.	0.	31.	0.	
. *	. *	179	075	191	000	

			0.	0.	33.	0.
. *		. .	131	030	580	000
			0.	-	34.	0.
. .		. .	069	0.021	252	000
			0.	-	34.	0.
. .		. .	044	0.013	526	000
			0.	0.	35.	0.
. .		. .	061	024	054	000
			0.	0.	35.	0.
. *		. .	077	040	906	000
			-	-	36.	0.
. .		* .	0	0.060	0.126	434 000
			0.	0.	36.	0.
. .		. *	1	048	102	772 000
			-	-	37.	0.
. .		* .	2	0.042	0.111	037 000
			-	-	38.	0.
* .		* .	3	0.112	0.083	897 000
			-	-	40.	0.
* .		. .	4	0.096	0.037	282 000
			-	-	45.	0.
* .		* .	5	0.181	0.142	248 000
			-	0.	46.	0.
* .		. .	6	0.086	057	375 000
			-	-	49.	0.
* .		* .	7	0.148	0.105	741 000
			-	-	55.	0.
* .		* .	8	0.196	0.079	692 000
			-	-	62.	0.
** .		* .	9	0.206	0.075	339 000
			-	-	69.	0.
** .		* .	0	0.212	0.124	430 000

Кореляція присутня тільки на 1 лазі при граничному значенні 0,2. Додатково можна перевірити на 3, 10, 15 лазі, якщо за граничне значення взяти 0,12. Результати оцінювання параметрів моделі АРКС(1, 1) представлені в таблиці 3.20 нижче:

Таблиця 3.20 Результати оцінювання моделі авторегресії з ковзним середнім першого порядку

Dependent Variable: Z
Method: ARMA Maximum Likelihood (BFGS)
Date: 05/07/20 Time: 21:41
Sample: 1 133
Included observations: 133
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
	44487.34	52684.60	0.844409	0.4000
C	0.9984	0.016289	61.29589	0.0000
AR(1)	0.4536	0.106044	4.277619	0.0000
MA(1)	2403056.	55880.97	43.00311	0.0000
SIGMASQ				
R-squared	0.995904	Mean dependent var		70547.61
Adjusted R-squared	0.995808	S.D. dependent var		24312.39
S.E. of regression	1574.030	Akaike info criterion		17.64103
Sum squared resid	3.20E+08	Schwarz criterion		17.72796
Log likelihood	-1169.129	Hannan-Quinn criter.		17.67636
F-statistic	10454.41	Durbin-Watson stat		1.725443
Prob(F-statistic)	0.000000			

Рівняння має вигляд:

$$y(k) = 44487,34 + 0,998y(k - 1) + mv(k) + 0,454mv(k - 1) \quad (3.17)$$

Основні статистичні характеристики моделі такі:

$$R^2 = 0,995; DW = 1,725; СКП = 3,20 \cdot 10^8.$$

Дана модель за статистикою Дарбіна-Уотсона, яка складає $DW = 1,725$, виходить досить точною. Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 1), які подані нижче на рисунку 3.17:

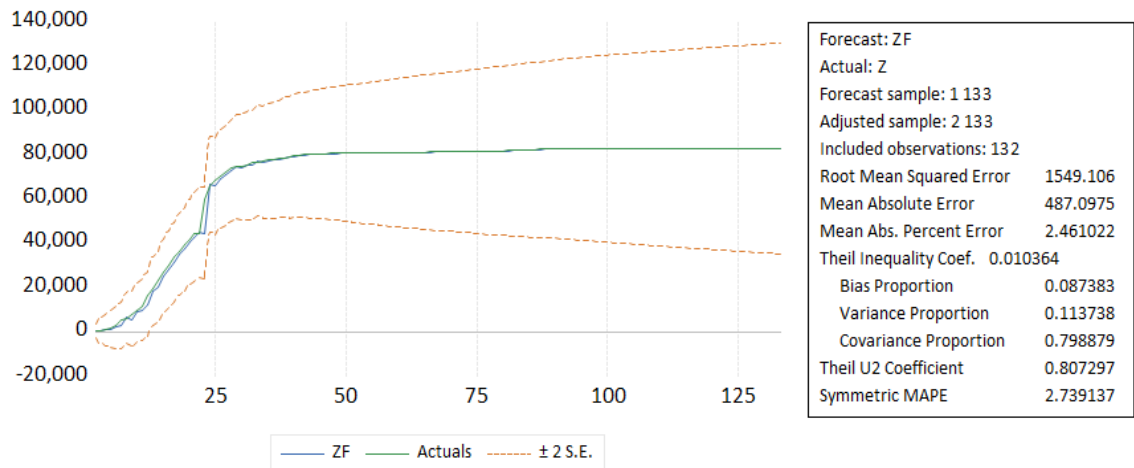


Рисунок 3.17 Характеристики однокрокового прогнозу за моделлю АРКС(1,1)

Таким чином, характеристики однокрокового прогнозу мають такі значення:

$$RMSE = 1549,11; \text{САП} = 487,098; \text{САПП} = 2,46; U = 0,0104.$$

За допомогою даної моделі отримали прогноз високої якості, тому що середня абсолютна похибка у процентах (САПП або MAPE) склала: САПП = 2,46 %. Це набагато краще результат і з точки погляду точності побудованої моделі, і якості прогнозу цієї моделі, ніж ми отримали при АР(1), АР(1) з квадратичним трендом.

Перевіримо для наочності моделі АРКС(1, q) та їх прогнози для q=3, 10, 15. Результати побудованих моделей зведено в одну таблицю (табл. 3.21), подану нижче.

Таблиця 3.21 Зведена таблиця для побудованих моделей АРКС(1,q) випадків захворювання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АРКС(1, 1)	0,995	$3,20 \cdot 10^8$	1,725	1549,1	487,098	2,46	0,0104
АРКС(1, 2)	0,996	$2,99 \cdot 10^8$	1,88	1499,0	450,5	2,25	0,01

АРКС(1, 3)	0,996	$2,85 \cdot 10^8$	1,9	1463,3	429,04	2,14	0,0098
АРКС(1, 10)	0,997	$2,42 \cdot 10^8$	1,99	1348,0	375,22	1,88	0,009
АРКС(1, 10) +КС(15)	0,997	$2,39 \cdot 10^8$	1,9995	1340,7	387,88	1,859	0,0089

Найкращими за показниками якості, а саме значення САПП = 1,859%, виявилась модель АРКС(1, 10)+КС(15), при чому з високим значенням статистики Дарвіна-Уотсона DW=1,9995. Параметри рівняння АРКС(1, 10)+КС(15):

$$\begin{aligned}
 y(k) = & 49451,33 + 0,995y(k-1) + mv(k) + 0,443mv(k-1) \\
 & + 0,199mv(k-2) + 0,241mv(k-3) + 0,307mv(k-4) \\
 & + 0,234mv(k-5) + 0,175mv(k-6) + 0,111mv(k-7) \\
 & + 0,107mv(k-8) + 0,224mv(k-9) + 0,003mv(k-10) \\
 & - 0,104mv(k-15)
 \end{aligned} \quad (3.18)$$

Розглянемо показники якості побудованого прогнозу моделі, які зображені нижче на рисунку 3.18:

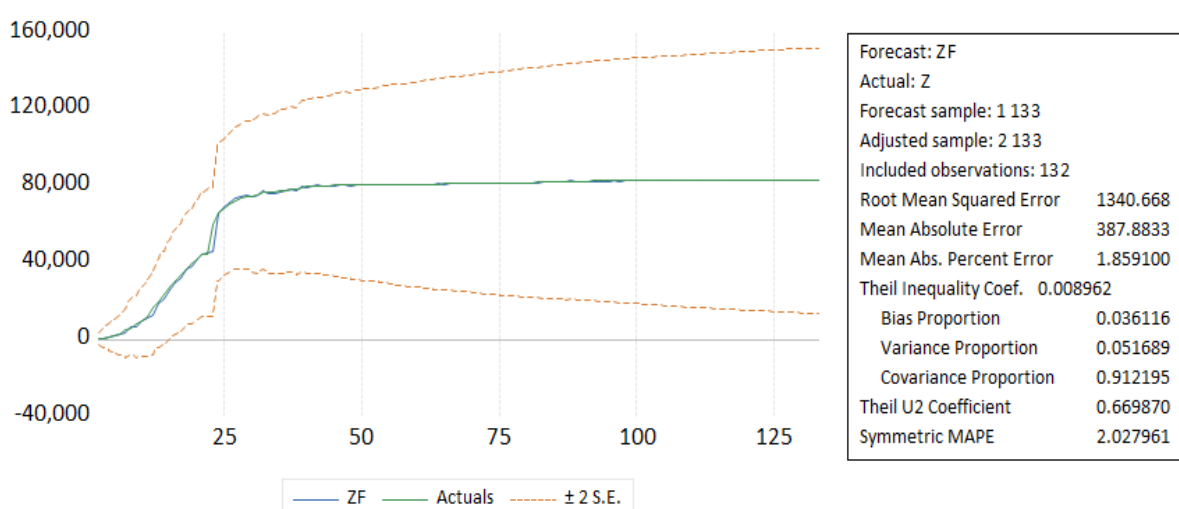


Рисунок 3.18 Характеристики однокрокового прогнозу за цією моделлю

Таблиця 3.20 наочно свідчить про суттєве підвищення якості однокрокового прогнозу за моделлю авторегресійного рівняння з ковзним середнім. Найкращою серед усіх розглянутих, з показником середня абсолютна похибка у відсотках (САПП або MAPE) САПП = 1,86%, модель АРКС(1,10)+КС(15).

3.2.4 Побудова моделей для прогнозування числа тих, хто одужав в Китаї

Як і в минулих прикладах, розглядається процес накопичення кількості жителів, які одужали від коронавірусу. Дані взято за період з 22.01.2020 по 02.06.2020, так як з кількістю зареєстрованих випадків, на момент початку ведення статистики уже було 28 осіб, що одужали. Графік цього процесу подано на рис. 3.19:

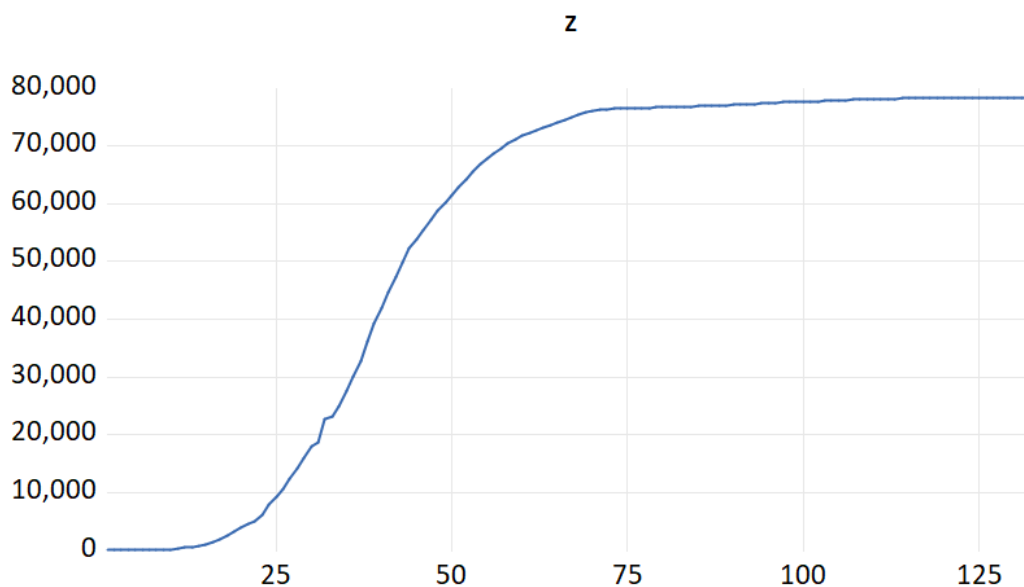


Рисунок 3.19 Кількість осіб, що одужали на корона-вірус в Китаї за період з 22.01.2020 по 02.06.2020

Графік характерний для процесу поширення епідемії, при чому для Китаю спалах епідемії давно пройшов свій пік і добігає кінцю згідно з графіком. Тож аналогічно попереднім прикладам спробуємо дослідити даний часовий ряд і побудувати моделі та прогнози для нього. Результати побудованих моделей зведено в одну таблицю (табл. 3.22), подану нижче.

Таблиця 3.22 Зведена таблиця для побудованих моделей випадків одужання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АР(1)	0,998	1,48*10 ⁸	0,2169	1055,3	599,24	5,24	0,00842
АР(1)+t3	0,999	33376975	0,909	502,85	364,32	66,48	0,004
АР(1)+t2	0,999	62707309	0,4775	689,24	543,83	39,07	0,0055
АРКС(1, 1)	0,999	79455162	1,1782	772,79	409,19	4,028	0,0062
АРКС(1, 2)	0,999	53131265	1,2552	631,93	304,41	3,417	0,005
АРКС(1, 3)	0,999	36548413	1,3267	525,14	257,79	3,14	0,0042
АРКС(1, 5)	0,999	26870760	1,7782	449,47	214,98	2,9924	0,0036
АРКС(1, 14)	0,999	13435473	1,9141	327,87	159,48	3,114	0,0026
АРКС(1, 5) +КС(14)	0,999	26870735	1,7783	449,47	214,96	2,9925	0,0027

Найкращими за показниками якості прогнозу, а саме значення САПП = 2,99%, виявилась модель АРКС(1, 5), при чому з досить високим значенням статистики Дарвіна-Уотсона DW=1,778. Параметри рівняння АРКС(1, 5):

$$\begin{aligned}
 y(k) = & 40092,21 + 0,999y(k-1) + mv(k) + 0,474mv(k-1) \\
 & + 0,868mv(k-2) + 0,738mv(k-3) + 0,446mv(k-4) \\
 & + 0,354mv(k-5)
 \end{aligned}
 \tag{3.20}$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 5), які зображені нижче на рисунку 3.20:

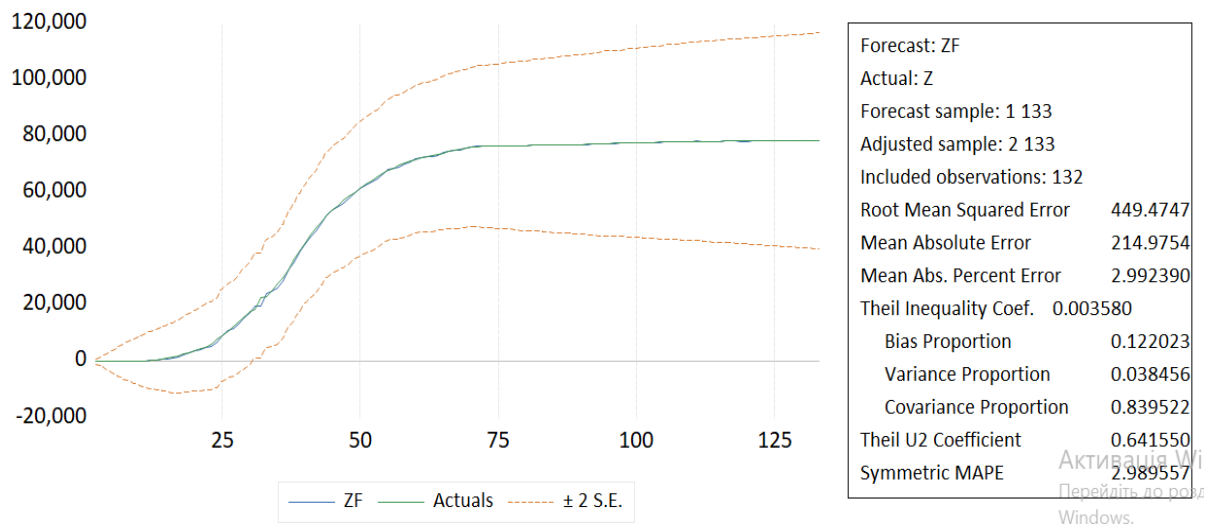


Рисунок 3.20 Характеристики однокрокового прогнозу за моделлю АРКС(1,5)

Отож, завдяки ковзному середньому вдалося значно покращити показники якості прогнозу моделі для часового ряду, який відображає кількість осіб, що одужали в період з 22.01.20 до 02.06.20 в Китаї.

3.2.5 Побудова моделей для прогнозування числа тих, хто захворів в Іспанії

В Іспанії дуже швидко поширився коронавірус, вся система охорони здоров'я країни пережила неймовірні перегрузки тому, що в порівнянні з Китаєм, якому вдалося локалізувати епідемію і цим мінімізувати наслідки, уряд Іспанії увів карантинні обмеження, коли епідемія уже набирала свої оберти, що миттєво сказалося на кількості осіб, які захворіли (більше 230 000

при населенні в 49,94 млн. чоловік, в порівняння населення Китаю – 1,393 млрд.)

Тож розглянемо процес накопичення кількості жителів, які захворіли на коронавірус. Дані по захворюваності взято за період з 01.02.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.21:

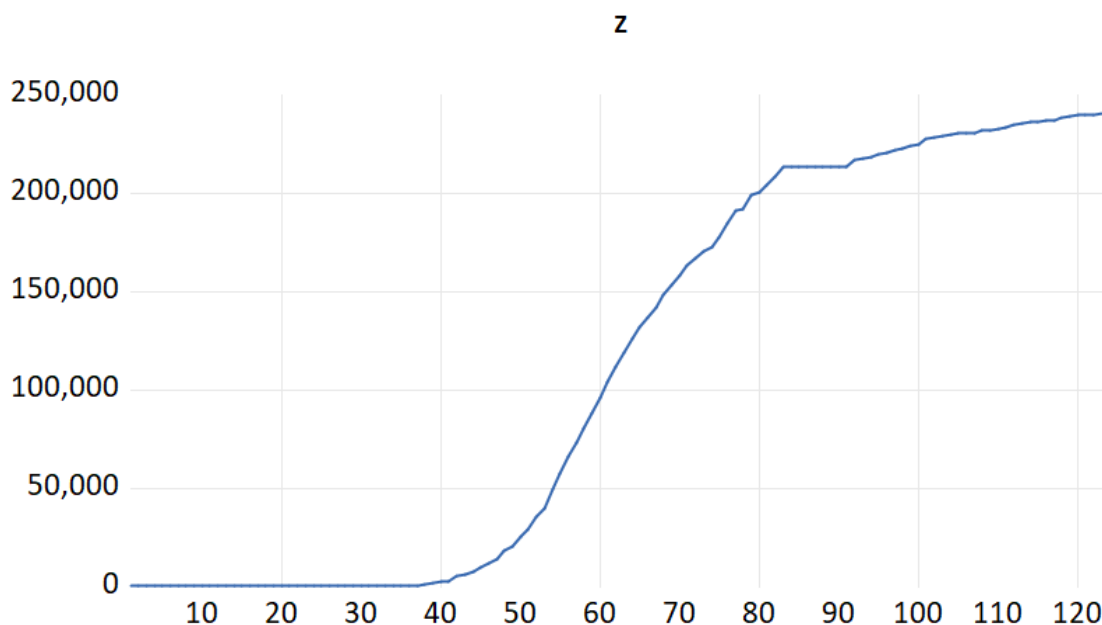


Рисунок 3.21 Захворюваність на корона-вірус в Іспанії за період з 01.02.2020 по 02.06.2020

У ситуації з Іспанією на радість дослідникам спостереження ведеться з початку епідемії (з першого офіційно виявленого випадку). На графіку промальовується логістична крива, яка пройшла свою точку перегину і наблизилась до своєї асимптоти. Про що це говорить нам, як дослідникам? Перша хвиля епідемії уже досягнула максимального значення нових випадків інфікування вдень і йде на спад. Не можна виключати ймовірності виникнення другої хвилі епідемії, поки не винайдена вакцина, як ефективний метод боротьби та протистояння виникненню епідемії.

Аналогічно попереднім прикладам побудуємо моделі та їхні прогнози для даного часового ряду. Результати побудованих моделей зведено в одну таблицю (табл. 3.23), подану нижче.

Таблиця 3.23 Зведена таблиця для побудованих моделей випадків захворювання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АР(1)	0,997	5,862	0,833	3569,32	2374,1	8,499	0,0119
АР(1)+t3	0,999	2,928	1,601	4350,94	2913,3	8,272	0,0145
АРКС(1, 1)	0,998	4,578	1,742	2781,41	1852,1	7,009	0,0093
АРКС(1, 2)	0,998	4,227	1,807	2481,35	1682,7	6,704	0,0083
АРКС(1, 3)	0,999	3,729	1,841	2281,86	1538,9	6,343	0,0076
АРКС(1, 15)	0,999	2,184	2,0	3504,03	2001,2	7,293	0,0117
АРКС(1, 16)	0,999	2,184	1,986	3624,78	2146,9	7,136	0,0121

Найкращими за показниками якості прогнозу, а саме значення САПП = 6,34%, виявилась модель АРКС(1, 3), при чому з досить високим значенням статистики Дарвіна-Уотсона $DW=1,841$. Параметри рівняння АРКС(1, 3):

$$y(k) = 6,428 + 0,998y(k-1) + mv(k) + 0,345mv(k-1) + 0,249mv(k-2) + 0,337mv(k-3) \quad (3.21)$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 3), які зображені нижче на рисунку 3.22:

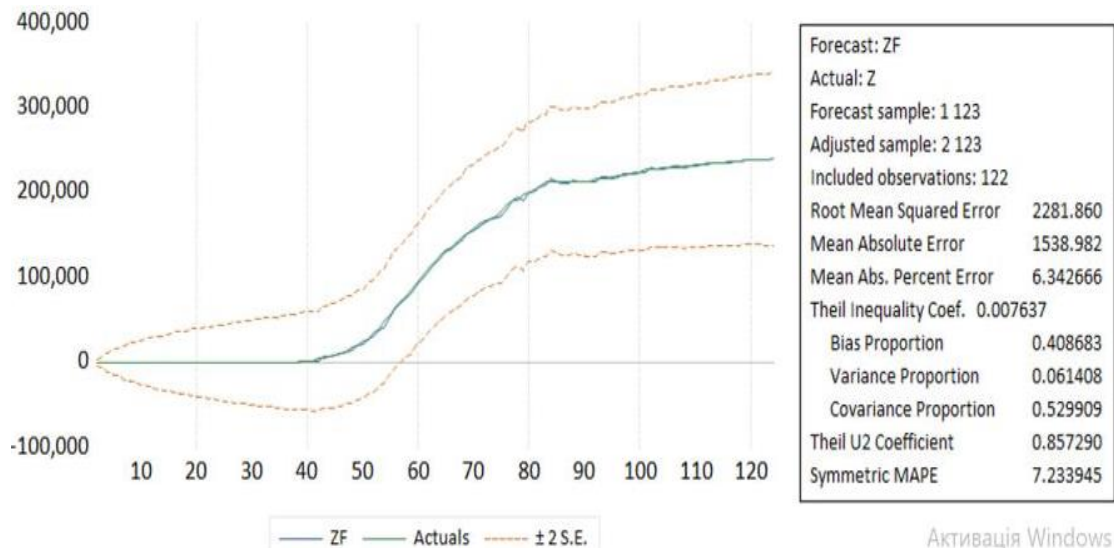


Рисунок 3.22 Характеристики однокрокового прогнозу за моделлю $APKC(1,3)$

В даному випадку завдяки ковзному середньому нам вдалося покращити показники якості прогнозу моделі для часового ряду, який відображає кількість осіб, що захворіли в період з 01.02.20 до 02.06.20 в Іспанії.

3.2.6 Побудова моделей для прогнозування числа тих, хто одужав в Іспанії

Як і в минулих прикладах, розглядається процес накопичення кількості жителів, які одужали від коронавірусу. Дані взято за період з 01.02.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.23:

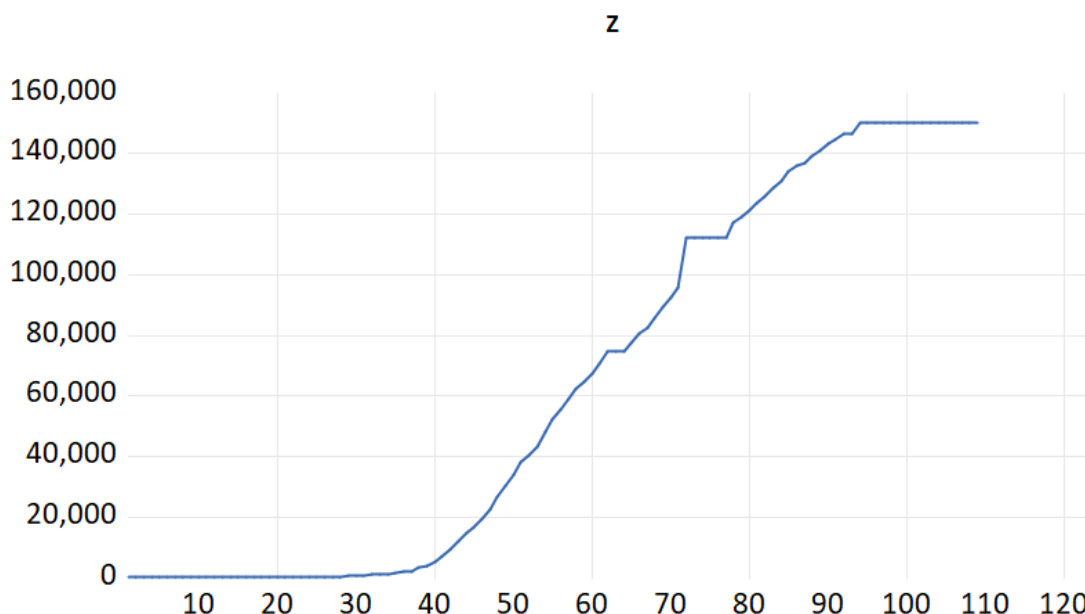


Рисунок 3.23 Кількість тих, хто одужав від корона-вірусу в Іспанії за період від 01.02.2020 до 02.06.2020

Результати побудованих моделей зведено в одну таблицю (табл. 3.24), подану нижче.

Таблиця 3.24 Зведена таблиця для побудованих моделей випадків одужання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АР(1)	0,999	12,942	1,863	3006,5	2021,8	7,774	0,0173
АР(1)+t1	0,994	11,185	2,141	2728,06	1800,4	11,047	0,0154
АРКС(1, 1)	0,994	12,896	1,983	2955,03	1980,3	7,916	0,017
АРКС(1, 2)	0,994	12,858	1,972	2930,56	1958,7	7,714	0,0169
АРКС(1, 3)	0,994	11,965	1,801	2899,85	1903,1	8,743	0,0167
АРКС(1, 4)	0,995	9,873	1,981	2757,17	1722,1	7,019	0,0159
АРКС(1, 6)	0,995	9,821	2,003	1973,66	982,039	570,95	0,01128
АРКС(1, 10)	0,999	3,29*10 ⁸	1,941	1737,99	919,793	662,95	0,0099

Найкращими за показниками якості прогнозу, а саме значення САПІ = 475,098%, виявилась модель АРКС(1, 2), при чому з досить високим значенням статистики Дарвіна-Уотсона $DW=1,833$, але отриманий результат не можна назвати навіть прийнятним, бо отриманий прогноз по цій моделі абсолютно неадекватний за своєю якістю.

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 2), які зображені нижче на рисунку 3.24:

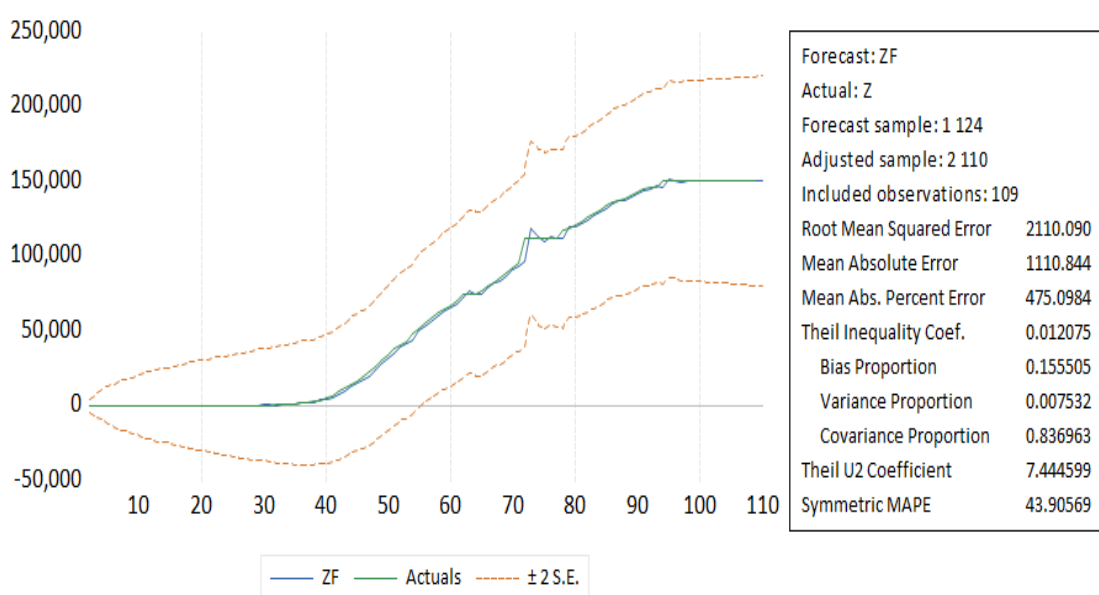


Рисунок 3.24 Характеристики однокрокового прогнозу за моделлю АРКС(1,2)

В даному випадку навіть завдяки ковзному середньому не вдалося значно покращити показники якості прогнозу моделі для часового ряду, який відображає кількість осіб, що одужали в період з 01.02.20 до 02.06.20 в Іспанії.

3.2.7 Побудова моделей для прогнозування числа тих, хто захворів в Бразилії

Бразилія вже декілька тижнів гордо займає друге місце серед країн, в яких виявлений коронавірус. 6 червня 2020 року президент Бразилії Жаїр Болсонару офіційно підтримав рішення уряду частково приховувати надалі дані стосовно кількості інфікованих в країні, тому модель об'єктивніше будувати до цього дня. Незважаючи на те, що перший випадок в Бразилії був зафіксований мінімум через 36 днів (початок фіксації офіційно підтверджених випадків інфікування Covid-n19 в Китаї), число заражень коронавірусом уже перебільшило 740 000 при населенні 209,5 млн. осіб. Таке розходження в числах, наприклад з Китаєм, об'єктивно можна аргументувати низьким рівнем громадської, соціальної відповідальності, недостатньою підтримкою від держави медицини, запізненим введенням карантинних мір (в недостатній кількості, масштабі).

Тож розглянемо процес накопичення кількості жителів, які захворіли на коронавірус. Дані по захворюваності взято за період з 26.02.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.25:

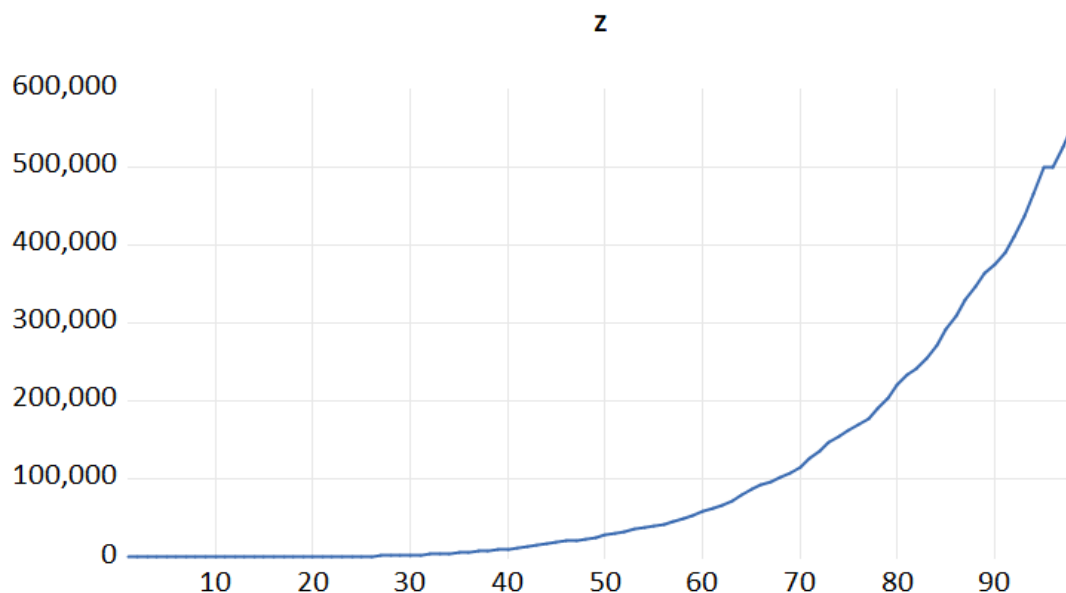


Рисунок 3.25 Захворюваність на корона-вірус в Бразилії за період з 01.02.2020 по 02.06.2020

Бразилія – це одна з країн, які ведуть статистику з першого підтвердженого випадку зараження.

Аналогічно попереднім прикладам побудуємо моделі та їхні прогнози для даного часового ряду, попередньо його прологарифмувавши. Результати побудованих моделей зведено в одну таблицю (табл. 3.25), подану нижче.

Таблиця 3.25 Зведена таблиця для побудованих моделей випадків захворювання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
AP(1)	0,996	5,4198	1,335	10284,1	6090,97	11,634	0,00293
AP(1)+t1	0,997	2,7971	2,512	3621,93	1723,15	7,946	0,0101
AP(1)+t2	0,998	2,7837	2,466	5662,34	2508,32	8,303	0,01587
АРКС(1, 1)	0,996	5,1199	1,816	9171,03	5374,07	11,191	0,02607
АРКС(1, 3)	0,997	4,0987	1,904	7395,87	4165,34	9,518	0,0209
АРКС(1, 4)	0,997	3,7628	2,017	6847,46	3776,5	8,94	0,0193
АРКС(1, 5)	0,997	3,736	2,019	6829,68	3743,27	9,29	0,01926
АРКС(1, 6)	0,998	2,716	1,833	5862,56	3301,26	8,313	0,0165
АРКС(1, 7)	0,998	2,4772	1,839	6545,6	3424,17	8,255	0,0184

Найкращими за показниками якості прогнозу, а саме значення САПП = 7,95%, виявилась модель AP(1)+t1, при чому з досить наближенням до ідеального значенням статистики Дарвіна-Уотсона DW=2,512. Параметри даного рівняння:

$$y(k) = 0,342 + 0,978y(k - 1) - 0,0002k \quad (3.23)$$

Розглянемо показники якості побудованого прогнозу моделі AP(1)+t1, які подані нижче на рисунку 3.26:

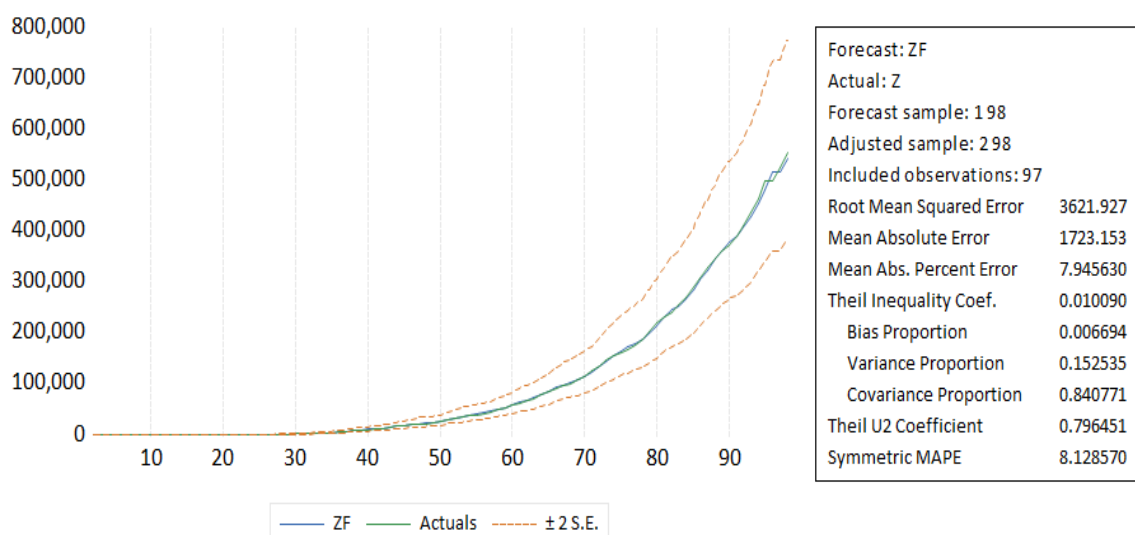


Рисунок 3.26 Характеристики однокрокового прогнозу за моделлю $AR(1)+t1$

В даному випадку завдяки моделі авторегресії першого порядку з трендом першого порядку вдалося побудувати досить якісний прогноз з середньою абсолютною похибкою у відсотках САПП=7,95% щодо кількості осіб, що захворіли в період з 26.02.20 до 02.06.20 в Бразилії.

3.2.8 Побудова моделей для прогнозування числа тих, хто одужав в Бразилії

Як і в минулих прикладах, розглядається процес накопичення кількості жителів, які одужали від коронавірусу. Дані взято за період з 17.03.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.27:

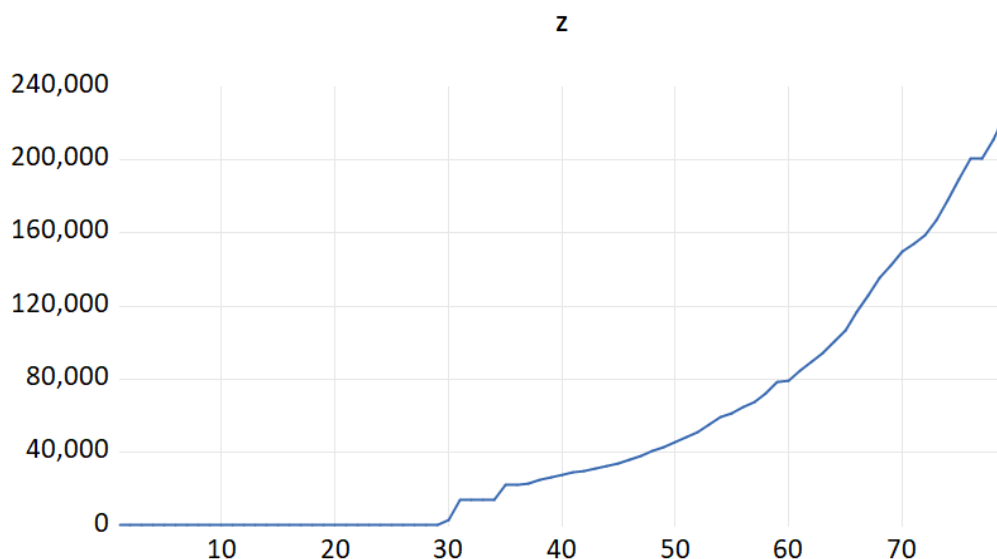


Рисунок 3.27 Кількість тих, хто одужав від корона-вірусу в Бразилії за період від 17.03.2020 до 02.06.2020

Результати побудованих моделей зведено в одну таблицю (табл. 3.26), подану нижче.

Таблиця 3.26 Зведена таблиця для побудованих моделей випадків одужання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
AP(1)	0,983	21,918	1,5937	0,526	0,173	4,394	0,0292
AP(1)+t1	0,985	18,767	1,7011	0,491	0,211	9,464	0,027
AP(1)+t2	0,987	18,324	1,7409	0,47	0,237	8,522	0,025
АРКС(1, 1)	0,984	20,839	2,014	0,513	0,181	5,028	0,0284
АРКС(1, 11)	0,983	22,773	1,8071	0,569	0,434	16,133	0,0316
АРКС(1, 14)	0,989	14,789	2,1423	0,5149	0,3308	12,117	0,0285
АРКС(1, 15)	0,991	12,4033	1,7057	0,467	0,226	5,439	0,0258
АРКС(1,1)+ КС(15)	0,989	14,634	1,7361	0,48	0,207	4,829	0,0266

АРКС(1, 17)	0,999	3,29*10 ⁸	1,9407	0,464	0,245	5,991	0,0256
-------------	-------	----------------------	--------	-------	-------	-------	--------

Найкращими за показниками якості прогнозу, а саме значення САПП = 4,39%, виявилась модель АР(1), при чому з допустимим значенням статистики Дарвіна-Уотсона $DW=1,5937$, при цьому отриманий результат не можна назвати найоптимальнішим. Таким варіантом стала модель АРКС(1, 1), яка має трохи більшу середню абсолютну похибку у відсотках САПП=5,03 (несуттєво відрізняється від значення моделі АР(1)) і майже ідеальне значення статистики Дарбіна-Уотсона $DW=2,014$. Параметри даного рівняння:

$$y(k) = 6,426 + 0,994y(k - 1) + mv(k) + 0,222mv(k - 1) \quad (3.24)$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 1), які подані нижче на рисунку 3.28:

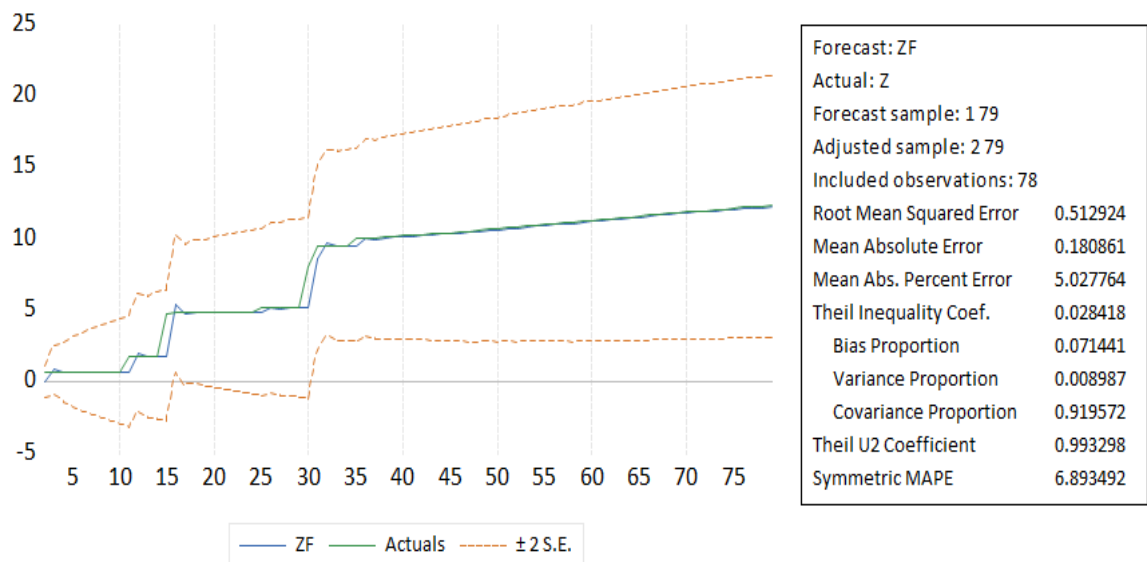


Рисунок 3.28 Характеристики однокрокового прогнозу за моделлю АРКС(1,1)

В даному випадку завдяки ковзному середньому вдалося побудувати модель авторегресії першого порядку з ковзним середнім першого порядку з

досить точним прогнозом динаміки числа осіб, що одужали в період з 17.03.20 до 02.06.20 в Бразилії.

3.2.9 Побудова моделей для прогнозування числа тих, хто захворів в США

Сполучені Штати Америки очолює рейтинг країн по кількості людей, які захворіли на коронавірус, більше того майже половина всіх зареєстрованих випадків зараження у світі виявлено на території США. Згідно статей американського видання The Washington Post з посиланнями на джерела, які ознайомлені зі звітами розвідувальних установ країни, президент ще у січні був проінформований про природу та ризики глобального поширення вірусу, про часткове приховання даних стосовно масштабів і серйозності проблеми урядом Китаю, про потенційну потребу запровадження урядових заходів для протидії поширенню коронавірусу і не сприйняв дане повідомлення серйозно на той момент. Хоча варто було, так як до країни щодня прибувало товаро-транспортними шляхами тисячі повітряних і морських суден з Європи та Китаю на пряму. Враховуючи інкубаційний період вірусу 2-3 тижні, він швидко поширився в головних транспортних вузлах Америки, а звідти і по всій країні, що і вивело США з експоненціальною швидкістю на перше місце в рейтингу.

Тож розглянемо процес накопичення кількості жителів, які захворіли на коронавірус. Дані по захворюваності взято за період з 24.01.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.29:

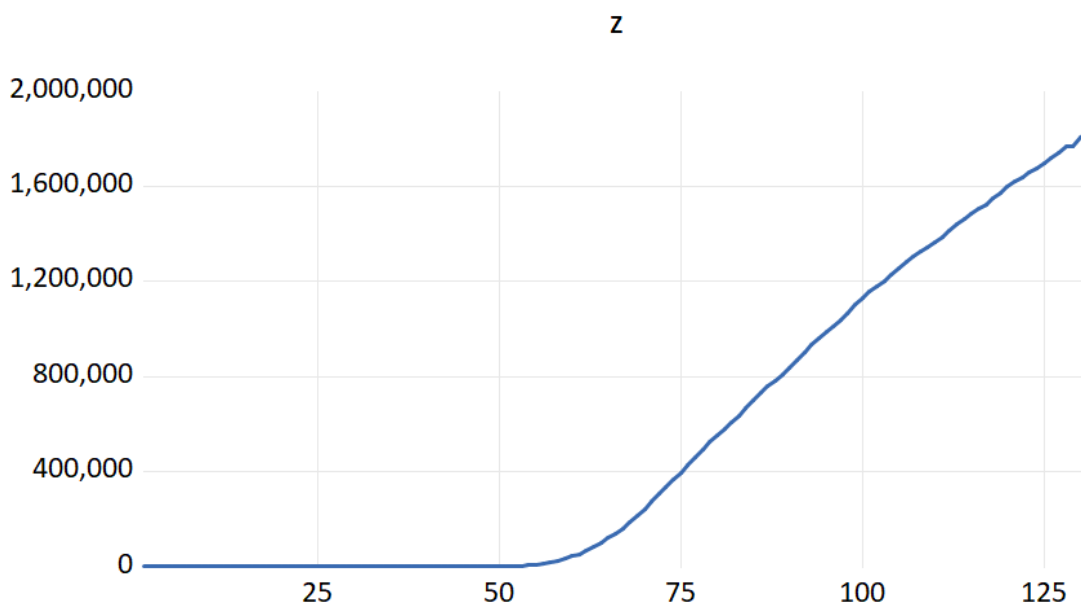


Рисунок 3.29 Захворюваність на корона-вірус в США за період з 24.01.2020 по 02.06.2020

Аналогічно попереднім прикладам побудуємо моделі та їхні прогнози для даного часового ряду, попередньо його прологарифмувавши. Результати побудованих моделей зведено в одну таблицю (табл. 3.27), подану нижче.

Таблиця 3.27 Зведена таблиця для побудованих моделей випадків захворювання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
AP(1)	0,998	4,493	0,992	20774,7	15282,1	9,226	0,0123
AP(1)+t1	0,999	2,672	1,673	23811,9	14867,9	8,937	0,0143
AP(1)+t2	0,999	2,432	1,785	32416,2	19212,1	8,358	0,01967
AP(1)+t3	0,999	2,395	1,844	16495,9	10470,7	7,535	0,01
АРКС(1, 1)	0,999	3,765	1,835	16088,9	11683,8	8,72	0,0098
АРКС(1, 2)	0,999	3,355	1,69	14163,0	10292,6	7,844	0,0087
АРКС(1, 3)	0,999	2,873	1,763	12135,8	8729,7	7,562	0,0074

АРКС(1, 5)	0,999	2,509	1,902	12537,8	7759,5	7,532	0,0076
АРКС(1, 8)	0,999	2,194	1,823	15855,7	9453,14	7,088	0,0097

Оптимальним варіантом виявилась модель АРКС(1, 8), яка показує якісний досить якісний прогноз з середньою абсолютною похибкою у відсотках САПП=7,09%, статистика Дарбіна -Уотсона при цьому DW=1,823. Модель АРКС(1, 8)+КС(26) дає незначне покращення, але різниця складає всього декілька сотих, тому раціональніше використати перший варіант. Параметри даного рівняння:

$$\begin{aligned}
 y(k) = & 7,056 + 0,999y(k-1) + mv(k) + 0,107mv(k-1) \\
 & + 0,385mv(k-2) + 0,599mv(k-3) + 0,319mv(k-4) \\
 & + 0,191mv(k-5) + 0,128mv(k-6) - 0,097mv(k-7) \\
 & + 0,395mv(k-8)
 \end{aligned} \quad (3.25)$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 8), які подані нижче на рисунку 3.30:

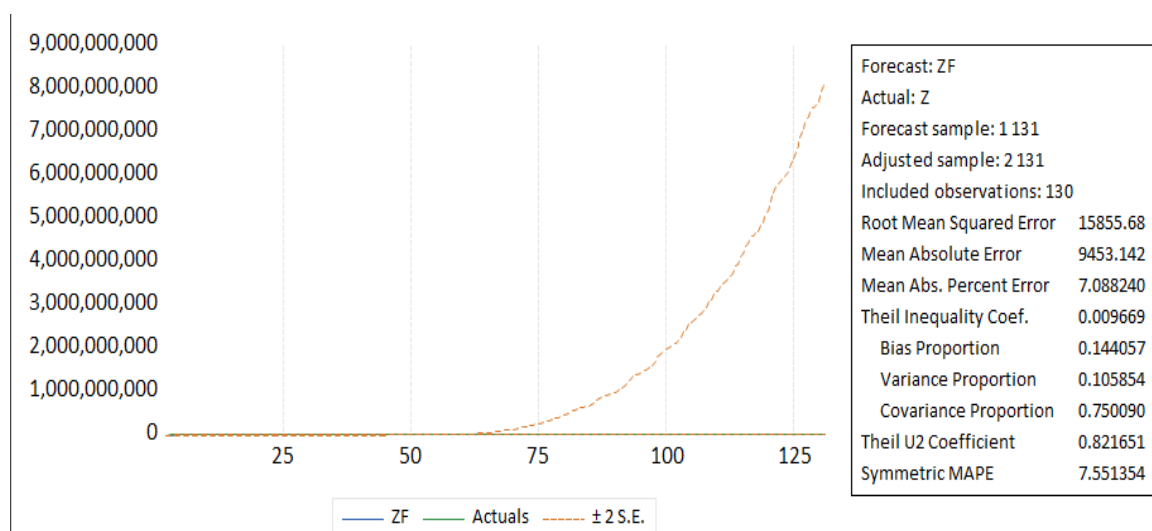


Рисунок 3.30 Характеристики однокрокового прогнозу за моделлю АРКС(1,8)

В даному випадку завдяки моделі авторегресії першого порядку з ковзним середнім восьмого порядку вдалося побудувати досить якісний прогноз з середньою абсолютною похибкою у відсотках САПП=7,09% щодо кількості осіб, що одужали в період з 24.01.20 до 02.06.20 в США.

3.2.10 Побудова моделей для прогнозування числа тих, хто одужав в США

Як і в минулих прикладах, розглядається процес накопичення кількості жителів, які одужали від коронавірусу. Дані взято за період з 09.02.2020 по 02.06.2020. Графік цього процесу подано на рис. 3.31:

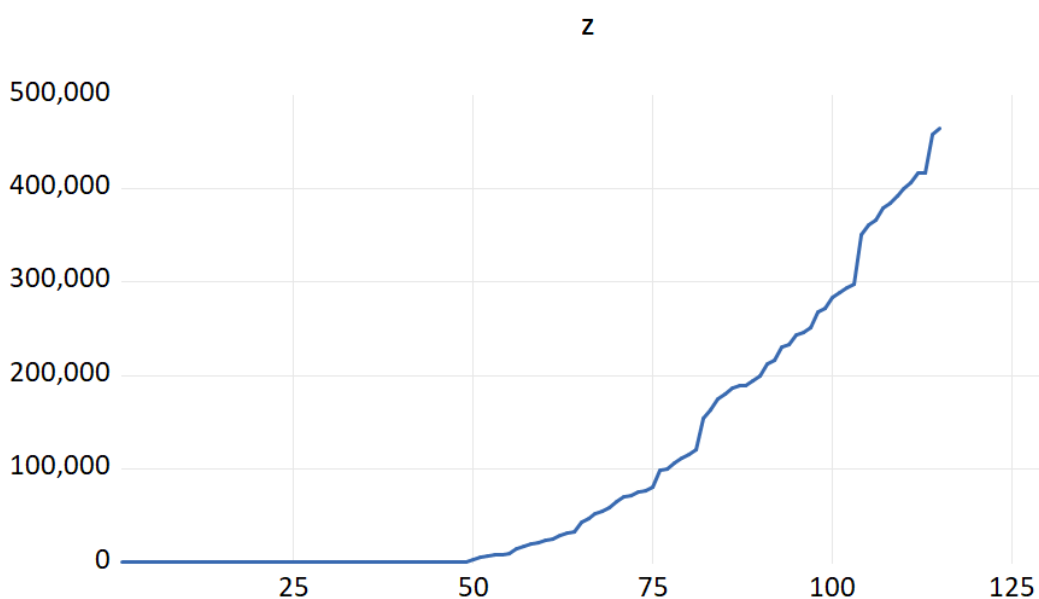


Рисунок 3.31 Кількість тих, хто одужав від корона-вірусу в США за період від 09.02.2020 до 02.06.2020

Результати побудованих моделей зведено в одну таблицю (табл. 3.28), подану нижче.

Таблиця 3.28 Зведена таблиця для побудованих моделей випадків одужання

Модель процесу	Характеристики моделі			Характеристики прогнозу			
	R^2	$\sum e^2$	DW	СКП	САП	САПП	Тейла
АР(1)	0,997	7,213	1,503	9119,8	4528,5	8,549	0,0277
АР(1)+t1	0,998	5,851	1,837	10383	5771,1	10,829	0,0302
АР(1)+t2	0,998	5,15	2,018	9246,0	4653,7	10,531	0,0277
АРКС(1, 1)	0,997	6,954	1,888	8791,9	4054,6	8,439	0,0267
АРКС(1, 2)	0,997	6,048	1,847	8500,2	3865,7	8,917	0,0257
АРКС(1, 6)	0,998	4,993	2,007	7649,3	3378,5	8,293	0,023
АРКС(1,6)+ КС(11)	0,998	4,822	1,995	7589,3	3352,6	8,096	0,0228
АРКС(1, 11)	0,998	4,623	1,885	7946,1	3651,1	7,967	0,0239

Найкращими за показниками якості прогнозу, а саме значення САПП = 7,97%, виявилась модель АРКС(1, 11), при чому з досить наближеним до ідеального значенням статистики Дарвіна-Уотсона $DW=1,885$. Параметри даного рівняння:

$$\begin{aligned}
 y(k) = & 7,257 + 0,997y(k-1) + mv(k) + 0,079mv(k-1) \\
 & + 0,337mv(k-2) + 0,068mv(k-3) + 0,036mv(k-4) \\
 & + 0,019mv(k-5) + 0,329mv(k-6) - 0,016mv(k-7) \\
 & + 0,208mv(k-8) + 0,094mv(k-9) - 0,0705mv(k-10) \\
 & + 0,204mv(k-11)
 \end{aligned} \tag{3.26}$$

Розглянемо показники якості побудованого прогнозу моделі АРКС(1, 11), які подані нижче на рисунку 3.32:

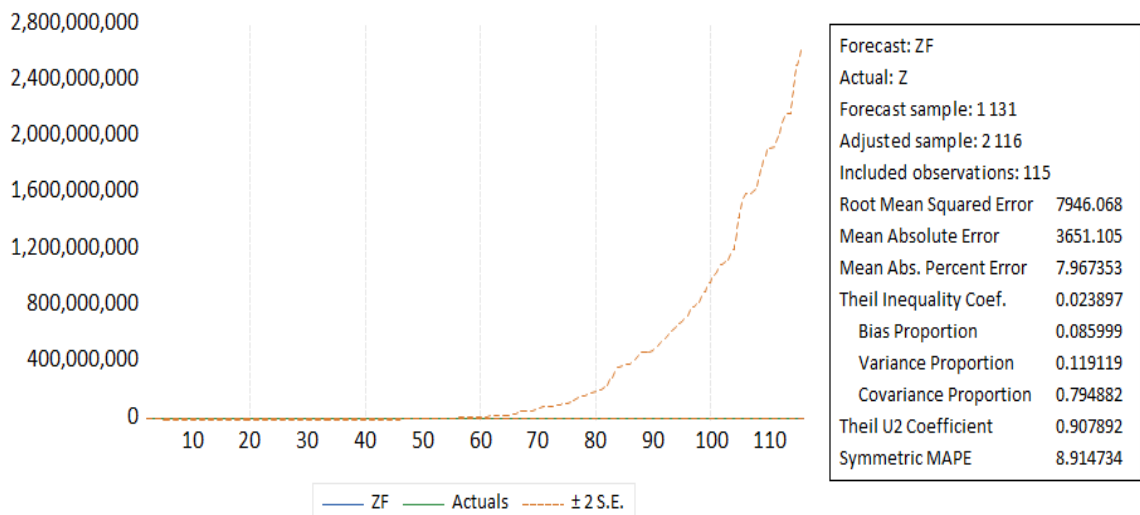


Рисунок 3.32 Характеристики однокрокового прогнозу за моделлю АРКС(1,1)

В даному випадку завдяки ковзному середньому вдалося побудувати модель авторегресії першого порядку з ковзним середнім 11-го порядку з досить точним прогнозом динаміки числа осіб, що одужали в період з 09.02.20 до 02.06.20 в США.

3.3 Альтернативна модель

Розглянемо модель: $V=I+R$, де V – загальна кількість постраждалих від інфекції; I – кількість тих, хто захворів і розповсюджує інфекцію; R – кількість тих, хто вже не розповсюджує інфекцію[1 – 3]. Сюди відносяться ті, хто знаходиться в ізоляції, всі, хто вже видужав, померлі та інфіковані, які вже залишили конкретну місцевість. Модель у диференціальній формі має вигляд:

$$\frac{dS}{dt} = -\alpha SI ; \quad (3.1)$$

$$\frac{dI}{dt} = \alpha SI - \rho I ; \quad (3.2)$$

$$\frac{dR}{dt} = \rho I, \quad (3.3)$$

де α і ρ – константи, які оцінюють на основі наявних статистичних даних.

Початкові умови для рівнянь (3.1) – (3.3) визначаються таким чином:

$$I(t_0)=1, R(t_0)=0, S(t_0)=N-1, N=S+I+R \quad (3.4)$$

Аналітичний розв’язок рівнянь (3.1) – (3.3) отримано шляхом введення функції $V(t) = I(t) + R(t)$, яка характеризує кумулятивну кількість підтверджених випадків захворювання [1, 2]:

$$F_1(V, N, \nu) = \alpha(t - t_0), \quad (3.5)$$

$$F_1 = \int_1^V \frac{dU}{(N-U)[\nu \ln(N-U) + U - \nu \ln(N-1)]}, \quad \nu = \frac{\rho}{\alpha}. \quad (3.6)$$

Таким чином, для кожної множини параметрів N, ν, α, t_0 і фіксованого значення V можна обчислити інтеграл (3.6) і знайти момент часу за виразом (3.5). Потім можна обчислити функції $I(t)$ і $R(t)$ за допомогою поданих нижче формул [1, 2]:

$$I = \nu \ln S - S + N - \nu \ln(N-1), \quad S = N - V, \quad R = V - I. \quad (3.7)$$

Функція I має максимум при $S = \nu$, і прямує до нуля на нескінченності. Кількість тих, хто підозрюється на нескінченності, $S_\infty > 0$, можна знайти за допомогою нелінійного рівняння [1, 2]:

$$S_\infty = (N-1)e^{\frac{S_\infty - N}{\nu}}. \quad (3.8)$$

Остаточну (фінальна) кількість жертв (остаточна кількість інфікованих) можна знайти за виразом:

$$V_{\infty} = N - S_{\infty}. \quad (3.9)$$

Для оцінювання тривалості епідемії можна скористатись умовою:

$$V(t_{final}) = 1, \quad (3.10)$$

Яка означає, що при $t > t_{final}$ можна не знайти ще одну особу, яка розповсюджує інфекцію. Очевидно, що подана вище модель має більш теоретичний ніж практичний характер, а тому розрахунки за цією моделлю не виконувались.

3.4 Моделювання процесу поширення пандемії в програмі AnyLogic

На даний момент в більшості країн, що борються з пандемією коронавірусу, введено карантинні міри різного ступеня обмеженості. Більше того, в зв'язку з економічною кризою, яку тянуть за собою введення обмежень на роботу багатьох підприємств, уряд вимушений розглядати поступовий план виходу всіх сфер діяльності і прошарків населення з карантину. З точки зору дослідження процесу поширення пандемії коронавірусу цікаво розглянути декілька варіантів розвитку подій і змоделювати різні сценарії симуляцій.

SIR-модель, на основі якої в програмі будуватимуться симуляції і яка була розглянута в другому розділі вище, дозволяє створити досить примітивні імітації взаємодії між собою особин популяції. Проте, проаналізувавши навіть

доволі умовні результати впливу різного роду факторів на життя популяції в спрощених імітаційних схемах, можна буде сформулювати висновки про те, що сприяє або пригнічує поширення вірусу в популяції.

В SIR-модель передбачено поділ популяції на три категорії:

- S-susceptible – ті, що можуть заразитися;
- I-infectious – ті, що вже інфіковані;
- R-recovered – ті, що одужали (точніше сказати ті, що вже не хворіють і під цим розуміти, що в цю категорію входять і ті, що одужали, і ті, що померли).

Для імітації фізичного контакту (обіймів, рукопотискань, чхання, дотику до різних поверхонь і т.д.), кожен з групи S, попадаючи в радіус зараження навколо кожного члена групи I, матиме шанс захворіти з ймовірністю p (початкове значення поставимо 20%) за кожен одиницю часу. Кожен інфікований через певний період часу одужує або помирає, тобто вибуває з числа тих, хто може поширювати вірус.

3.4.1 Моделювання базової симуляції

1). В першому варіанті розглянемо ситуацію, в якій ніхто не реагуватиме на поширення вірусу, тобто в симуляції точки рухаються просто в хаотичному порядку. Динаміку процесу поширення зображено на рисунку 3.33:

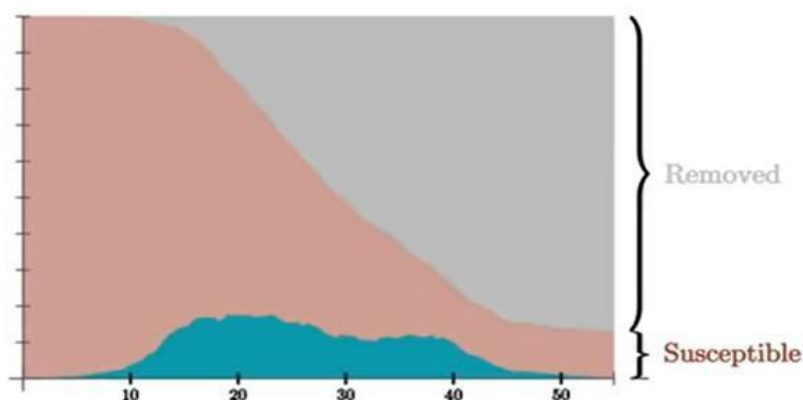


Рисунок 3.33 Графік поширення вірусу

R лежить в межах $[0, 2.22]$. Залишився невеликий відсоток осіб, які не постраждали.

Епідемія розвивалась природнім шляхом, згідно графіку фактор росту не піднімався вище 2.22, що збігається з фактичними даними по COVID-19 у світі. Для порівняння R_0 для COVID-19 приблизно 2, для іспанки в 1918 році теж 2, для сезонного грипу – 1,3.

2). Перевіримо логіку побудованої симуляції, збільшивши радіус зараження вдвічі. Тобто, нехай дано більш соціально активне суспільство, де спостерігається більша кількість взаємодій з оточуючими людьми. Динаміку поширення зображено на рисунку 3.34:

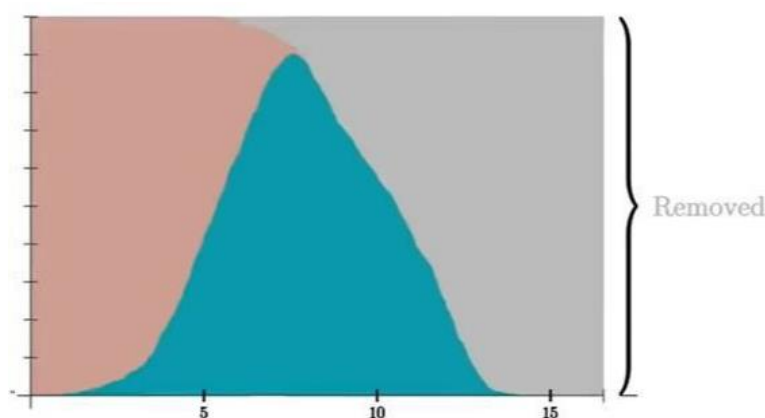


Рисунок 3.34 Графік поширення вірусу при збільшенні радіуса зараження вдвічі

Як і очікувалося збільшення радіуса зараження вдвічі посприяло набагато швидшому поширенню вірусу за короткий проміжок часу, при цьому варто помітити, що не залишилося осіб, які не постраждали на відмінно від першого випадку. А значення R лежить в межах $[0, 10]$.

3). Ще одна перевірка на логіку для змодельованої симуляції: повернемо початковий радіус зараження і зменшимо ймовірність зараження з 20% до 10%. Це імітує ситуацію, коли члени популяції стали ретельніше дотримуватися правил гігієни (частіше мити руки з милом, прикривати рот при кашлю, чханні, рідше торкатися руками лиця). Динаміку поширення зображено на рисунку 3.35:

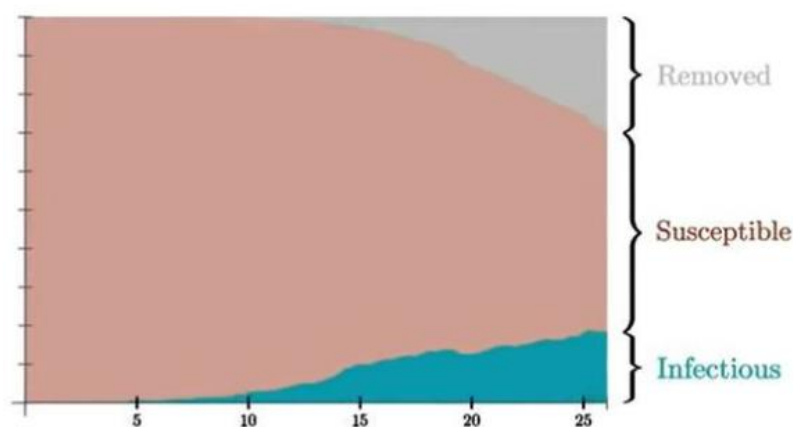


Рисунок 3.35 Графік поширення вірусу при зменшенні ймовірності зараження вдвічі

Зменшення ймовірності зараження сильно згладжує криву, тобто суттєво впливає на темпи поширення вірусу. Значний відсоток людей не постраждав. А значення R лежить в межах $[0, 1.62]$.

Як висновок, можна сказати, що вдалося наглядно продемонструвати наскільки фактор росту захворювання чутливий до кожного з контрольованих в першому варіанті моделі параметрів, а саме щоденна взаємодія, ймовірність зараження та тривалість хвороби. Для кожної окремо людини нескладно

зменшити за потреби кількість контактів за день і стежити за своєю гігієною, але в масштабах популяції це суттєво зменшить швидкість поширення вірусу, що в наслідку зачепить меншу кількість людей.

Адже, головною цілю в боротьбі з пандемією стоїть зменшення загальної долі осіб, що померли. Тобто не можна допускати ситуації, коли в один момент буде одночасно інфікована певна велика кількість людей (пік кривої інфікованих занадто високо на графіку, наприклад ситуація на рис. 3.25). Оскільки, в такому разі перевантажуються система охорони здоров'я та сильно зменшується доступність їхніх ресурсів, як результат помирає більший відсоток людей.

4). Одним з методів боротьби являється карантин або виявлення та ізоляція випадків зараження. Нехай в симуляції створимо карантинну зону, яка відкриється після виявлення певної кількості випадків інфікування (поставимо після 50 випадків). Введемо умову, що усі виявлені інфіковані особи через день відправлятимуться в карантин. Динаміка поширення зображена на рисунку 3.36:

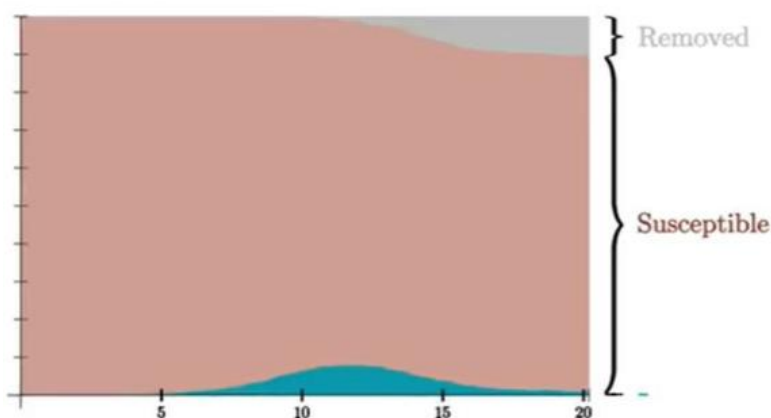


Рисунок 3.36 Графік поширення вірусу при умові ізоляції 100% хворих

R лежить в межах $[0, 1.82]$. Епідемія швидко зупинена. Проте, потрібно розуміти, що це в ідеальних умовах масштабного щоденного тестування та

ефективного виявлення можна ізолювати 100% інфікованих, але не в реаліях життя.

Тому введемо ймовірність виявлення випадків зараження 80%, наприклад у 20% захворювання протікає в латентній формі, тобто без симптомів. Динаміка поширення зображена на рис. 3.37:

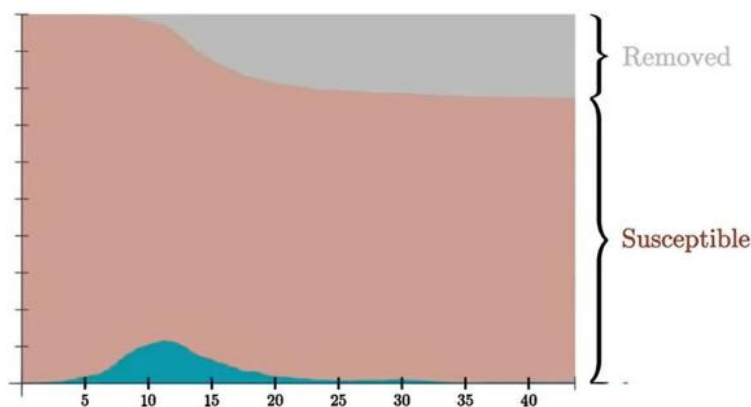


Рисунок 3.37 Графік поширення вірусу при умові ізоляції 80% інфікованих

R лежить в межах $[0, 1.82]$. Пік кривої інфікованих трохи більший, як у випадку зі 100%, і через невиявлені 20% хвіст кривої інфікованих набагато довший і протяжніший. Загалом випадків зараження виявилось в два рази більше, як у випадку з ізоляцією 100% інфікованих.

Як висновок, кількість людей, яка уникає тестування та перевірок, непропорційно сильно впливає на загальну кількість інфікованих.

З цієї причини, легко пояснити чому вірусні захворювання з вищою летальністю менш небезпечні в глобальному сенсі. Наприклад, вірус атипової пневмонії SARS (летальність близько 11%, для порівняння у його прямого нащадка COVID-19 3, 4%) був виявлений у 2002 році. Він поширився на 35 країн і всього захворіло 8461 особа, з яких померло 916, і майже всі мали симптоми хвороби, тому так швидко вдалося справитися з епідемією. Тобто через латентну або легку форму протікання захворювання, його стає важче ідентифікувати, що сприяє зараженню набагато більшої кількості людей, як

наслідок кількісно більше число смертей хоч ніби і від менш летальної хвороби.

5). Тепер спробуємо застосувати соціальне дистанціювання. Введемо умову включення коефіцієнта соціального дистанціювання після виявлення перших 50 випадків зараження. Ідеального дотримання правил соціального дистанціювання вимагати не можна так, як в симуляції обмежена зона існування точок-членів популяції і час від часу попадають в радіус зараження один одного, а в житті середньостатистичній людині неможливо повністю соціально дистанціюватися від решти. Проте, сенс в тому, що відбувається це набагато рідше.

Розглянемо динаміку поширення вірусу при 100%, 90%, 70%, 50% соц. дистанціюванні, представлену на рис. 3.38:

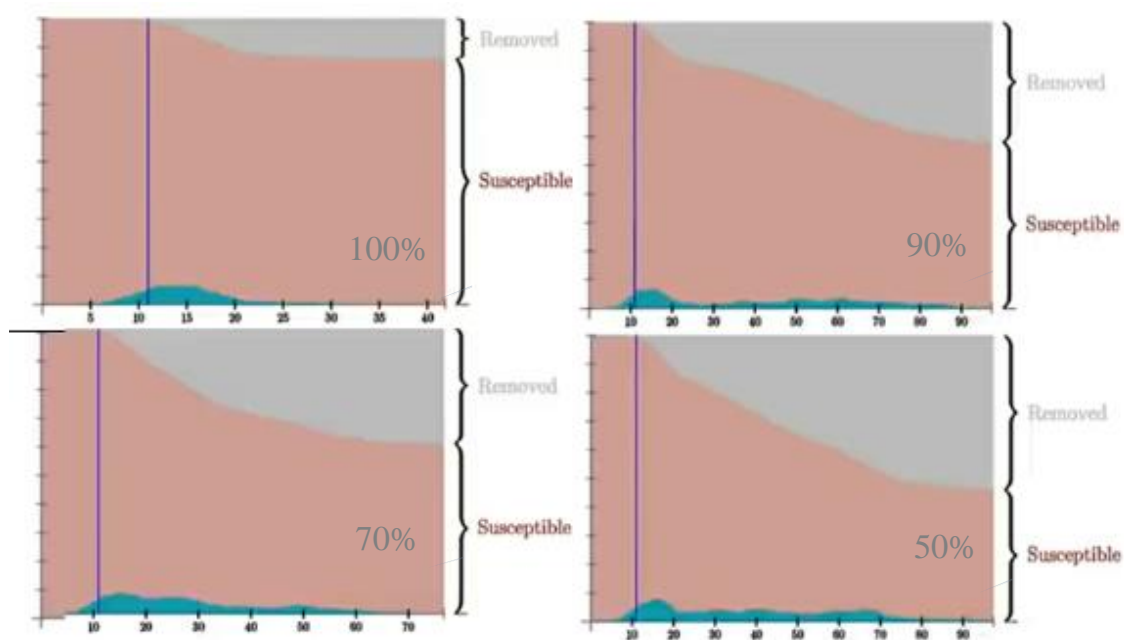


Рисунок 3.38 Графіки поширення вірусу за умов соціального дистанціювання 100%, 90%, 70%, 50% населення популяції

R лежить в межах $[0, 1.5]$. У всіх чотирьох варіантах застосування соціального дистанціювання сильно згладжує криву інфікованих, що

виключає можливість перевантаження системи охорони здоров'я та її ресурсів. Однак, звертаючи увагу на графіках на загальне число тих, хто заразився, то у варіантах 90%, 70% це трохи менше половини популяції, а у варіанті 50% чуть гірше - близько половини населення. Тобто насправді виявляється неважливим 10%, 30% чи 50% не дотримуються соціального дистанціювання. Якщо знайдуться хоча б 10% населення, які порушують встановлені правила, вони вноситимуть достатньо нестабільності і цього вистачить щоб хвіст кривої інфікованих ще довго тянувся.

В заключенні соціальне дистанціювання ідеально підходить для зглажування кривої, проте навіть незначні порушення такого режиму сприяють подальшому поширенню вірусу, негасання епідемії.

3.4.2 Моделювання симуляції з центральною точкою відвідування

1). Для підвищення достовірності результатів, створимо центральну точку відвідування для популяції, як імітацію реального магазину, школи, аптеки, робочого офісу, метро чи просто центра міста. Розглянемо динаміку поширення епідемії вірусу, зображену на рис. 3.39:

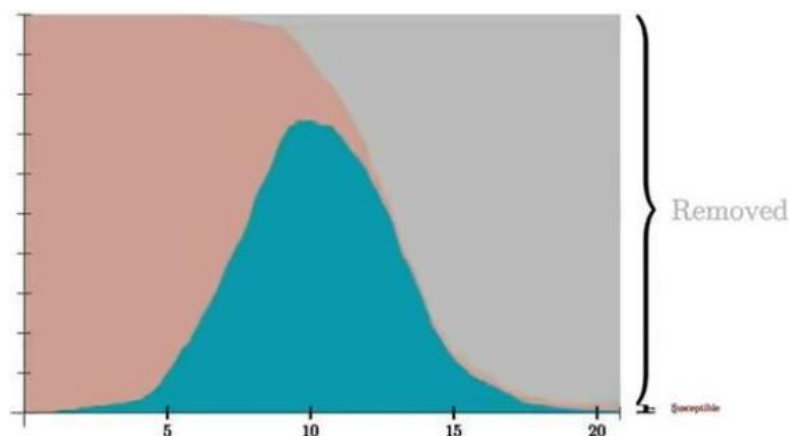


Рисунок 3.39 Графік поширення вірусу за наявності центру відвідування

R лежить в межах $[0, 5.8]$. Майже не залишилося осіб, які не постраждали. Отримані результати свідчать про те, що наявність центру, який регулярно відвідують люди сильно збільшує швидкість поширення вірусу та створює схожий ефект на ситуацію з більш соціально активною популяцією без центра. Цей графік буде контрольним для порівняння результатів, отриманих в цій симуляції, але під впливом різних факторів.

2). Для наближення симуляції до правдоподібності, врахуємо момент, що ймовірність зараження характеризує фізичний контакт (рукопотискання, обійми і т.д.), але характер взаємодії осіб в центральній точці дещо відрізняється від комунікації зі знайомими тобі людьми. Тому, зменшимо ймовірність зараження до 10%, тобто впововину. Динаміка поширення виглядатиме так, див. рис. 3.40:

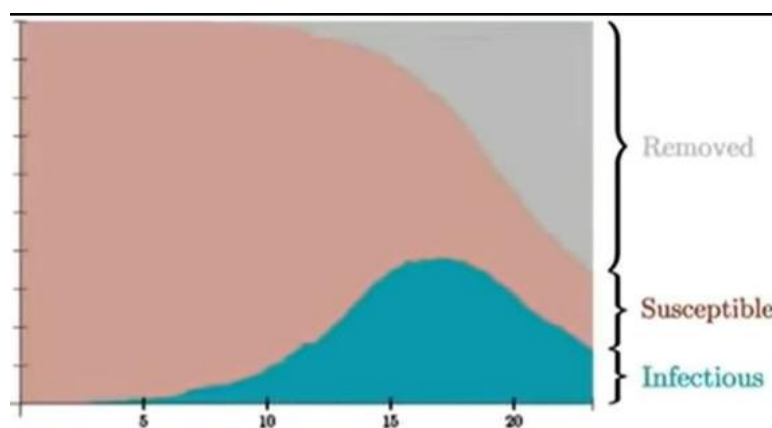


Рисунок 3.40 Графік поширення вірусу зі зменшенням ймовірності зараження вдвоє

Значення піку кривої інфікованих зменшилось вдвоє, проте ефект від наявності центральної точки відвідування не менший, епідемія доволі швидко поширилась і захворіло більше половини населення.

Тому висновки такі, що центральні точки відвідування, а точніше їх наявність значно збільшують швидкість поширення вірусу в будь-якому випадку.

3). Введемо соціальне дистанціювання після виявлення нехай 30 випадків захворювання. Наскільки це повпливає на кінцевий результат, розглянемо на графіку, рисунок 3.341

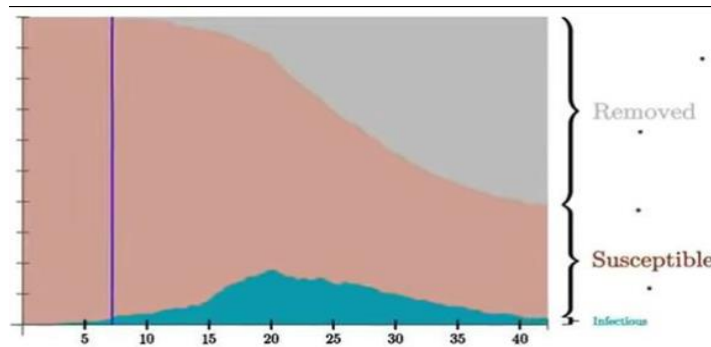


Рисунок 3.41 Графік поширення вірусу зі зменшеною ймовірністю зараження і дотриманням соціального дистанціювання

R лежить в межах $[0, 1.9]$. Пікове значення кривої інфікованих менше звичайно контрольного. Проте, загалом кінцевий результат майже однаковий чисельно. Тож, можна зробити висновки, що продовження відвідування центральної точки абсолютно нівелює введення соціальної дистанції, і навіть її дотримання.

Для перевірки цієї гіпотези поміняємо умову в симуляції і відключимо відвідування центральної точки при введенні соціальної дистанції. Динаміка поширення вірусу відображається на рисунку 3.42:

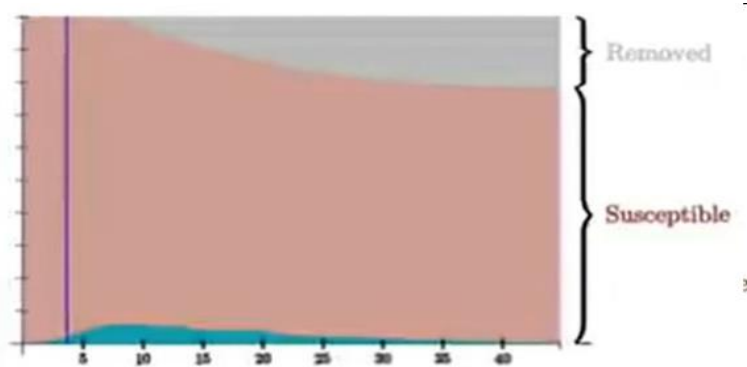


Рисунок 3.42 Графік поширення вірусу при введенні соціальної дистанції і закритті центру відвідувань

Тож з графіку помітно наскільки зменшилось загальне число інфікованих, тому гіпотезу підтвердили.

4). Для наглядності і порівняння результатів розглянемо два шляхи розвитку подій боротьби з ефектом центра відвідувань, яку можна вести:

1. Зменшення в 5 разів частоти відвідувань центральної точки до значення 0.01, але це сильно вплине на звички і побут людей.
2. Зменшення ще вдвоє ймовірності зараження до значення 0.05, тобто це означає ретельне стеження за власною гігієною.

Простежимо динаміку поширення вірусу для цих сценаріїв на рис.3.43:

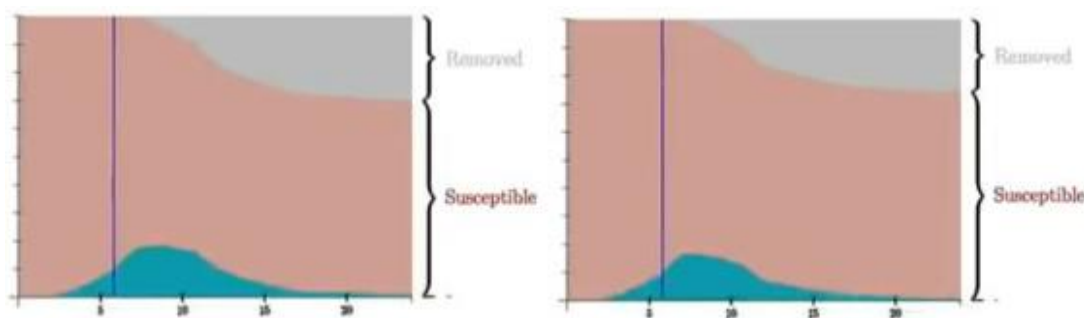


Рисунок 3.43 Графіки поширення вірусу за 1. та 2. сценаріями

Як можна побачити з графіків, вони майже ідентичні, тобто результати при застосуванні будь-якого з двох сценаріїв будуть майже однакові. Висновок такий, що ефективна гігієна має велике значення і вплив на результати поширення вірусу.

5). Усі популярні методи боротьби, які на даний момент використовуються, перевірялися окремо для оцінки та аналізу їхньої ефективності. Прослідкуємо динаміку поширення вірусу, коли примініться і соціальне дистанціювання (3.0), і зменшення частоти відвідувань центральної точки (0.01), і зменшення ймовірності зараження (0.05) на рисунку 3.44:

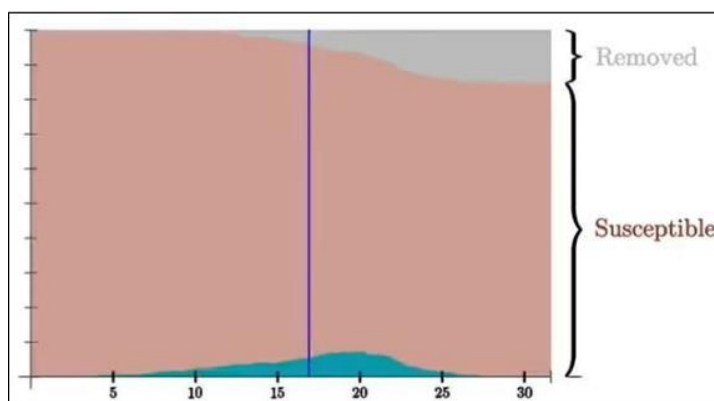


Рисунок 3.44 Графік поширення вірусу при введенні карантинних заходів

R лежить в межах $[0, 1.9]$. Отриманий результат говорить про досить високу ефективність комбінації розглянутих методів боротьби з пандемією коронавірусу у світі. Проте, потрібно пам'ятати, що таких результатів добилися в майже ідеальних лабораторних умовах, тобто через примітивну симуляцію процесу поширення вірусу в програмі. \

Зрештою, найкращим варіантом в боротьбі з поширенням вірусів є ефективна система тестування, виявлення та ізоляції підтверджених випадків зараження. Такий підхід гарантує швидкість подолання проблеми без впливу на якість цього результату рівня соціальної відповідальності та громадянської дисципліни.

ВИСНОВКИ ДО РОЗДІЛУ

Проаналізувавши результати моделювання на основі наявних статистичних даних, бачимо, що не всі моделі дали хороший результат, але в основному вдалось знайти модель, яка забезпечує високу якість короткострокового прогнозу розповсюдження епідемії. Побудовано дві регресійні моделі з трендами (тобто для нестационарних процесів), які дають

можливість спрогнозувати кумулятивну кількість інфікованих в Україні та кількість тих, хто одужає. При цьому досягнуто такі характеристики прогнозів: САПП = 15,37% і САПП = 8,84%.

Розглянута альтернативна модель у формі диференціальних рівнянь також корисна, але має більш теоретичне застосування. Тому було використано SIR-модель для побудови системи імітаційного моделювання і для отримання порівняльних результатів. Перевірено було популярні методи боротьби з поширенням вірусу, які використовуються на даний момент. Проаналізовано, як впливають централізовані точки на швидкість поширення епідемії та чому так необхідно застосування карантинних заходів в комплексі.

РОЗДІЛ 4: ФУНКЦІОНАЛЬНО-ВАРТІСНИЙ АНАЛІЗ ПРОГРАМНОГО ПРОДУКТУ

4.1 Постановка задачі проектування програмного продукту для моделювання та прогнозу поширення коронавірусу

Необхідно оцінити можливі варіанти реалізації власного програмного продукту для моделювання процесів поширення коронавірусу та оцінки їх прогнозів. Оцінка буде проводитися за допомогою функціонально-вартісного аналізу.

4.2 Обґрунтування функцій програмного продукту

Головна функція F_0 – розробка програмного продукту для моделювання та прогнозування процесів поширення коронавірусу. Виходячи з конкретної мети, можна виділити наступні основні функції ПП:

- F_1 – вибір мови програмування:
 - А. мова програмування C++;
 - Б. мова програмування Python;
- F_2 – спосіб введення вихідних даних;
 - А. зчитування з файлу;
 - Б. ініціалізація даних в кодї програми;
 - В. введення даних вручну за допомогою інтерфейсу програми.
- F_3 – інтерфейс користувача;
 - А. віконний інтерфейс;
 - Б. консольний інтерфейс.

- F₄ – виведення результатів роботи програми.
 - А. на екран з додатковою візуалізацією;
 - Б. в окремий файл.

Побудуємо морфологічну карту варіантів реалізації базових основних функцій, що буде відображати усі можливі варіанти реалізації основних функцій на рис.4.1:

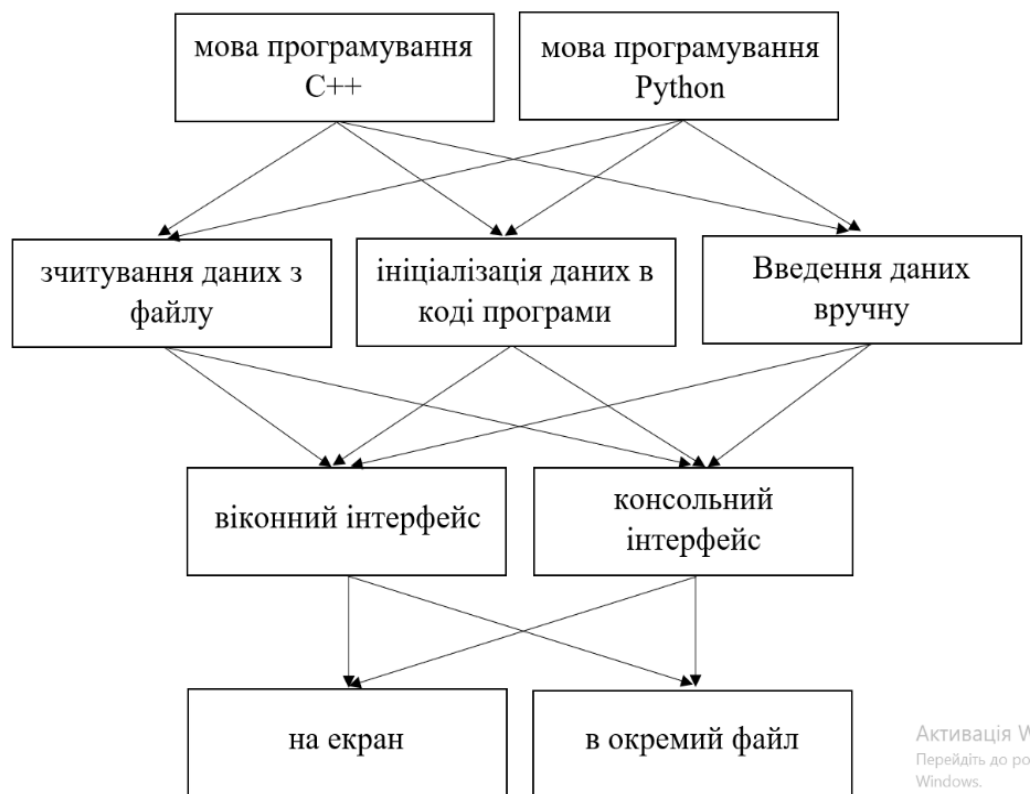


Рисунок 4.1 - Морфологічна карта.

Спираючись на карту була побудована матриця переваг та недоліків (табл. 4.1)

Таблиця 4.1 - Позитивно-негативна матриця

Основна функція	Варіант реалізації	Переваги	Недоліки
F1	А	Швидкодія	Ускладнена структура коду, багато часу на його написання

	Б	Простота реалізації, зручність	Невисока швидкодія
F2	А	Простота реалізації, ергономічність	Вимоги до формату файла
	Б	Легкість редагування	Зміна усіх займає велику кількість часу
	В	Не вимагає реалізації	Високі витрати часу на введення даних
F3	А	Не потребує додаткового ПЗ	Велика к-ть часу на реалізацію
	Б	Не потребує реалізації	Низька ергономічність, відсутність візуалізації
F4	А	Збереження даних	Низька ергономічність
	Б	Висока ергономічність	Проблематичне збереження даних

Для характеристики прототипу програмного додатку використовуємо параметри X1 – X5(табл. 4.2):

Таблиця 4.2 - Система параметрів додатку

Найменування параметру	Позначення параметру	Значення параметру		
		Гірше	Середнє	Краще
Швидкодія програмного коду, Оп/мс	X1	35000	50000	70000
Об'єм оперативної пам'яті, Мб	X2	40	25	15
Час обробки даних, мс	X3	1500	800	400
Потенційний об'єм коду, к-сть рядків	X4	800	600	400

На рисунках 4.2 – 4.5 зображено графіки значень параметрів додатку:

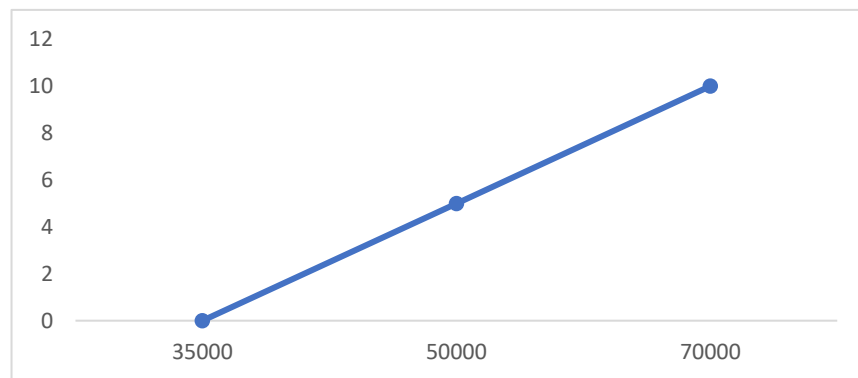


Рисунок 4.2 - Швидкодія програмного коду

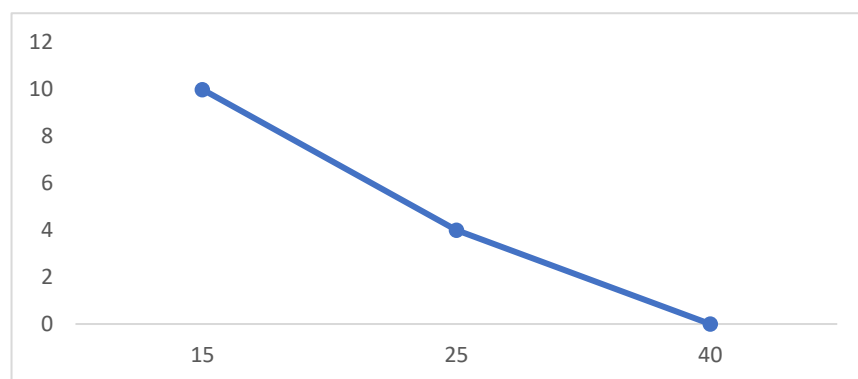


Рисунок 4.3 - Об'єм оперативної пам'яті

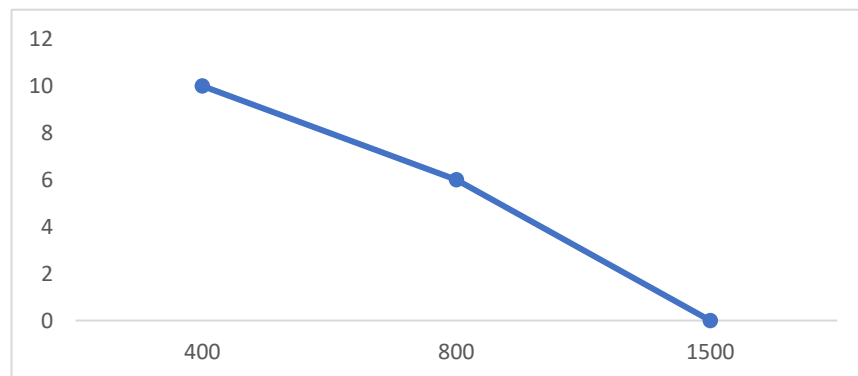


Рисунок 4.4 - Час обробки даних

Тепер попарно порівняємо наші параметри, зобразимо результати у вигляді таблиці 4.3:

Таблиця 4.3 - Попарне зрівняння параметрів

Параметри	Експерти							Кінцева оцінка	Числове значення
	1	2	3	4	5	6	7		
X1 і X2	<	<	<	<	<	<	<	<	0,5
X1 і X3	<	<	<	<	<	<	<	<	0,5
X1 і X4	>	>	<	>	>	>	<	>	1,5
X2 і X3	<	<	>	<	>	<	<	<	0,5
X2 і X4	>	>	>	>	>	>	>	>	1,5
X3 і X4	>	>	>	>	>	>	>	>	1,5

Далі необхідно обчислити коефіцієнт конкординації, скористаємося такою формулою:

$$W = \frac{12S}{N^2(n^3-n)} = \frac{12 \cdot 205}{7^2(4^3-4)} = 0,84 \quad (4.1)$$

Отже, в нашому випадку коефіцієнт конкординації дорівнює 0,84, що задовольняє вимогам до його значення (повинен бути більше, ніж 2/3). Далі продемонстровано табл. 4.4:

Таблиця 4.4 - Розрахунок вагомості параметрів

Параметри x_i	Параметри x_j				Перший крок		Другий крок		Третій крок	
	X1	X2	X3	X4	b_i	K_{bi}	b_i^1	K_{bi}^1	b_i^2	K_{bi}^2
X1	1,0	0,5	0,5	1,5	4,5	0,219	26,25	0,216	110	0,215
X2	1,5	1,0	0,5	1,5	5,0	0,281	29,25	0,282	134	0,283
X3	1,5	1,5	1,0	1,5	6,0	0,344	37,25	0,347	159,5	0,348
X4	0,5	0,5	0,5	1,0	3,5	0,156	22,25	0,155	84,5	0,154
Всього:					19	1	115	1	488	1

Отже, буде розглянуто такі варіанти реалізації програмного продукту:

1. $F1Б - F2A - F3A - F4A$
2. $F1Б - F2A - F3A - F4Б$

Далі представлено табл. 4.6 рівня якості альтернатив основних функцій:

Таблиця 4.6 – Рівень якості альтернатив основних функцій

Основна функція	Варіант реалізації	Абсолютне значення параметру	Бальна оцінка параметра	Коефіцієнт вагомості параметру	Коефіцієнт якості
$F1(X1)$	A	60000	6,1	0,215	1,3115
$F2(X2)$	A	20	5,0	0,283	1,415
$F3(X3)$	B	500	8,3	0,348	2,8884
$F4(X4)$	A	700	2,6	0,154	0,4004
	B	450	8,7	0,154	1,3398

Обрахуємо коефіцієнти якості кожного з варіантів розробки:

$$K_{K1} = 1,3115 + 1,415 + 2,8884 + 0,4004 = 6,0153$$

$$K_{K2} = 1,3115 + 1,415 + 2,8884 + 1,3398 = 6,9547$$

За розрахованими значеннями бачимо, що кращим є другий варіант, оскільки $6,9547 > 6,0153$.

4.3 Економічний аналіз варіантів розробки

Для цього необхідно оцінити трудомісткість розробки ПП, що включає такі завдання:

- 1) Зчитування та перевірка коректності даних

2) Розробка побудови моделі для прогнозування

3) Написання алгоритму обробки та виведення даних в файл

Це для першого варіанту реалізації, а для другого останнє завдання таке:

4) Надання можливості перегляду опису моделі та результатів роботи у вікні

Завдання 1-4 за ступенем новизни відноситься до групи В. За складністю алгоритми, які використовуються в завданні 2, 3 належать до групи 1; а в завданні 1, 4 – до групи 3.

Для другого та третього завдання, виходячи із норм часу для завдань розрахункового характеру ступеню новизни В та групи складності алгоритму 1, трудомісткість дорівнює: $T_p = 43$ людино-днів, а для 1,4 завдання з групою складності алгоритму 3, то $T_p = 12$. Поправочний коефіцієнт, який враховує вид нормативно-довідкової інформації для другого завдання: $K_{\Pi} = 1.35$. Поправочний коефіцієнт, який враховує складність контролю вхідної та вихідної інформації для всіх системи завдань рівний 1: $K_{СК} = 1$. При розробці використовуються стандартні модулі, врахуємо це - $K_{СТ} = 0.8$.

Повна трудомісткість 2, 3 завдання(складність – 1, новизна – В):

$$T_{2,3} = 43 \cdot 1.35 \cdot 0.8 = 46,44 \text{ людино-днів.}$$

Для 1, 4 завдання (використовується алгоритм третьої групи складності, ступінь новизни В), $K_{\Pi} = 1, K_{СК} = 1, K_{СТ} = 0.8$:

$$T_{1,4} = 12 \cdot 1 \cdot 0.8 = 9,6 \text{ людино-днів.}$$

Складаємо трудомісткість відповідних завдань для кожного з обраних варіантів реалізації програми, щоб отримати їх трудомісткість:

$$T_I = (9,6 + 46,44 + 46,44) \cdot 8 = 819,84 \text{ людино-годин;}$$

$$T_{II} = (9,6 + 46,44 + 9,6) \cdot 8 = 525,12 \text{ людино-годин};$$

Варіант II вимагає менших працевитрат.

Далі вважається, що робочий день складає 8 годин, в тиждні п'ять робочих днів. В розробці бере участь один програміст з окладом 18000 грн, один аналітик з окладом 20000 та тестувальник з окладом 15000 грн. Визначимо середню заробітну плату за годину:

$$C_{\text{ч}} = \frac{18000 + 20000 + 15000}{3 \cdot 22 \cdot 8} = 100,38$$

Тоді, розрахуємо витрати на виплату заробітної плати розробникам за варіантами:

$$\text{I. } C_{\text{ЗП}} = 100,38 \cdot 819,84 \cdot 1,2 = 98\,754,65 \text{ грн.}$$

$$\text{II. } C_{\text{ЗП}} = 100,38 \cdot 525,12 \cdot 1,2 = 63\,253,85 \text{ грн.}$$

Відрахування на соціальне страхування(22%)(грн):

$$\text{I. } C_{\text{ВД}} = C_{\text{ЗП}} \cdot 0,3677 = 98\,754,65 \cdot 0,22 = 21\,726,02 \text{ грн.}$$

$$\text{II. } C_{\text{ВД}} = C_{\text{ЗП}} \cdot 0,3677 = 63\,253,85 \cdot 0,22 = 13\,915,85 \text{ грн.}$$

Далі розрахуємо витрати на оплату однієї машино-години. Так як одна ЕОМ обслуговує одного програміста з окладом 18000 грн., з коефіцієнтом зайнятості 0,2 то для однієї машини отримаємо:

$$C_{\text{Г}} = 12 \cdot M \cdot K_3 = 12 \cdot 18000 \cdot 0,2 = 43200 \text{ грн.}$$

З урахуванням додаткової заробітної плати:

$$C_{3П} = C_{Г} \cdot (1 + K_3) = 43200 \cdot (1 + 0.2) = 51840 \text{ грн.}$$

Відрахування на єдиний соціальний внесок:

$$C_{ВІД} = C_{3П} \cdot 0.3677 = 51840 \cdot 0.22 = 11404,8 \text{ грн.}$$

Амортизаційні відрахування розраховуємо при амортизації 20% та вартості ЕОМ – 10000 грн.

$$C_A = K_{ТМ} \cdot K_A \cdot Ц_{ПР} = 1.15 \cdot 0.2 \cdot 10000 = 2300 \text{ грн.,}$$

Витрати на ремонт та профілактику розраховуємо як:

$$C_P = K_{ТМ} \cdot Ц_{ПР} \cdot K_P = 1.15 \cdot 10000 \cdot 0.05 = 575 \text{ грн.,}$$

Ефективний годинний фонд часу ЕОМ за рік розраховуємо за формулою:

$$T_{ЕФ} = (Д_K - Д_В - Д_С - Д_Р) \cdot t_3 \cdot K_В = (365 - 104 - 8 - 16) \cdot 8 \cdot 0.9 = 1706.4 \text{ годин,}$$

Витрати на оплату електроенергії розраховуємо за формулою:

$$C_{ЕЛ} = T_{ЕФ} \cdot N_C \cdot K_3 \cdot Ц_{ЕН} = 1706,4 \cdot 0,32 \cdot 0,3 \cdot 1.55 = 255.55 \text{ грн.,}$$

Накладні витрати розраховуємо за формулою:

$$C_H = Ц_{ПР} \cdot 0.67 = 10000 \cdot 0,67 = 6700 \text{ грн.}$$

Тоді, річні експлуатаційні витрати складатимуть:

$$C_{\text{ЕКС}} = 51840 + 11404,8 + 2300 + 575 + 255,55 + 6700 = 73075,75 \text{ грн.}$$

Собівартість однієї машино-години ЕОМ дорівнюватиме:

$$C_{\text{М-Г}} = C_{\text{ЕКС}} / T_{\text{ЕФ}} = 73075,75 / 1706,4 = 42,8 \text{ грн/год.}$$

В рамках реалізації даного програмного продукту ведуться на ЕОМ, в такому випадку витрати на оплату машинного часу будуть розраховуватися за наступною формулою:

$$C_{\text{М}} = C_{\text{М-Г}} \cdot T \quad (4.2)$$

$$\text{I. } C_{\text{М}} = 42,8 \cdot 819,84 = 35\,089,15 \text{ грн.};$$

$$\text{II. } C_{\text{М}} = 42,8 \cdot 525,12 = 22\,475,14 \text{ грн.};$$

$$C_{\text{Н}} = C_{\text{ЗП}} \cdot 0,67$$

$$\text{I. } C_{\text{Н}} = 98\,754,65 \cdot 0,67 = 66\,165,62 \text{ грн.};$$

$$\text{II. } C_{\text{Н}} = 63\,253,85 \cdot 0,67 = 42\,380,08 \text{ грн.};$$

Таким чином, витрати на розробку програмного продукту будуть складати:

$$C_{\text{ПП}} = C_{\text{ЗП}} + C_{\text{ВІД}} + C_{\text{М}} + C_{\text{Н}} \quad (4.3)$$

$$\text{I. } C_{\text{ПП}} = 98\,754,65 + 21\,726,02 + 35\,089,15 + 66\,165,62 = 221\,735,44 \text{ грн.};$$

$$\text{II. } C_{\text{ПП}} = 63\,253,85 + 13\,915,85 + 22\,475,14 + 42\,380,08 = 142\,024,92 \text{ грн.};$$

4.4. Вибір кращого варіанта ПП техніко-економічного рівня

Для визначення кращої альтернативи з техніко-економічної точки зору, розрахуємо коефіцієнти:

$$K_{K1} = 1,3115 + 1,415 + 2,8884 + 0,4004 = 6,0153$$

$$K_{K2} = 1,3115 + 1,415 + 2,8884 + 1,3398 = 6,9547$$

$$K_{\text{ТЕР}1} = 6,0153 / 221\,735,44 = 2,7128 \cdot 10^{-5};$$

$$K_{\text{ТЕР}2} = 6,9547 / 142\,024,92 = 4,8968 \cdot 10^{-5};$$

Таким чином альтернатива 1 є більш ефективною з техніко-економічної точки зору.

ВИСНОВКИ ДО РОЗДІЛУ

В цьому розділі було проведено функціонально-вартісний аналіз програмного продукту, що розробляється в рамках дипломного проектування. Процес аналізу складався з двох етапів.

На першому етапі були визначені основні функції продукту та альтернативні варіанти їх реалізації, а для визначення оптимальної альтернативи з технічної точки зору обчислювався рівень якості кожної з них.

На другому етапі виконувалося дослідження альтернатив на оптимальність з економічної точки зору, тут виконувалися обрахування витрат на трудові ресурси та матеріальне забезпечення, що необхідне для реалізації

програмного продукту та визначався коефіцієнт ефективності для визначення кращої з альтернатив.

Таким чином, за результатом проведеного двоетапного аналізу, було визначено, що кращою є друга альтернатива, отже обрано такий варіант реалізації програмного продукту:

- Мова програмування – Python;
- Зчитування даних з файлу;
- Віконний інтерфейс;
- Виведення результатів на екран з додатковою візуалізацією.

ВИСНОВКИ ДО РОБОТИ

По мірі прогресування пандемії COVID-19 країни все частіше приймають найрізноманітніші заходи. Наші результати показують, що необхідно буде накласти безліч мір, незалежно від того, чи є пригнічення або пом'якшення наслідків основною політичною метою. Однак пригнічення вимагатиме напашування більш інтенсивних і соціально деструктивних заходів, ніж пом'якшення наслідків. Вибір заходів в кінцевому підсумку залежить від відносної доцільності їх здійснення і їх можливу ефективність в різних соціальних контекстах.

Відносна ефективність різних заходів, що відрізняються від досвіду країн на сьогоднішній день, є складним завданням, оскільки багато хто з них здійснили кілька (або всі) з цих заходів з різним ступенем успіху. За допомогою госпіталізації всіх випадків (а не тільки тих, які вимагають стаціонарного лікування), Китай фактично ініціював форму ізоляції випадків, скоротивши передачу інфекції від випадку до випадку в домашньому господарстві і в інших умовах. У той же час, завдяки впровадженню соціальної дистанції в масштабах всього населення, можливість передачі інфекції в подальшому в усіх місцях була швидко зменшена. За оцінками декількох досліджень, ці заходи дозволили знизити R до рівня нижче 1. Останнім часом ці заходи почали слабшати. Пильне спостереження за ситуацією в Китаї останні два місяці повинне посприяти розробці стратегій в інших країнах.

В цілому, отримані в 3 розділі результати показують, що соціальне відчуження в масштабах всього населення, що застосовується до населення в цілому, матиме найбільший вплив і в поєднанні з іншими заходами, зокрема, ізоляцією будинків і закриттям шкіл і університетів - потенційно може придушити передачу інфекції нижче порога $R = 1$, необхідного для швидкого зниження захворюваності. Таким чином, мінімальною політикою ефективного пригнічення є соціальне відчуження всього населення в поєднанні з

домашньою ізоляцією випадків зараження і закриттям шкіл і університетів. Щоб уникнути поновлення передачі вірусу, цю політику необхідно буде підтримувати до тих пір, поки не з'являться великі запаси вакцини для імунізації населення, що може скласти 18 і більше місяців.

Однак існують дуже великі невизначеності щодо передачі цього вірусу, можливої ефективності різних стратегій і ступеня, в якому населення мимовільно приймає ризиковані моделі поведінки. Це означає, що важко точно визначити ймовірну первісну тривалість заходів, які будуть потрібні, за винятком того, що вона складе кілька місяців. Майбутні рішення про те, коли і протягом якого періоду часу послаблювати політику, повинні будуть прийматися на основі постійного спостереження.

Заходи, які використовуються для досягнення цілі ослаблення карантину, можуть також еволюціонувати з плином часу. У міру зменшення числа випадків стає все більш доцільним приймати інтенсивні випробування, контактне відстеження та карантинні заходи, аналогічні стратегіям, які використовують сьогодні в Південній Кореї. Такі технології, як додатки для мобільних телефонів, які відстежують взаємодію людини з іншими людьми в суспільстві, можуть дозволити зробити таку політику більш ефективною і популярною, якщо вдасться подолати пов'язані з цим проблеми, пов'язані з недоторканністю приватного життя. Проте, якщо інтенсивні пакети нефармацевтичних заходів, спрямовані на придушення, що не будуть збережені, наш аналіз показує, що передача інфекції швидко відновиться, що потенційно може призвести до епідемії, порівнянної за масштабами з тією, яку можна було б спостерігати, якби не було вжито жодних заходів втручання.

Поєднання ізоляції виявлених випадків захворювання, карантину домогосподарств і соціальної дистанційованості осіб з більш високим ризиком важких наслідків (літніх людей та осіб з іншими базовими захворюваннями) є найбільш ефективним поєднанням політичних заходів для пом'якшення наслідків епідемії. Ізоляція випадків захворювання, і карантин домогосподарств є основними епідеміологічними заходами щодо пом'якшення

наслідків інфекційних захворювань і спрямовані на зниження можливості подальшої передачі інфекції шляхом зниження рівня контакту з тими, хто, як відомо, є інфекційним (випадки зараження) або може бути носієм інфекції (контакти з домогосподарствами).

Потрібно відмітити, що немає ніякої впевненості в тому, що заходи з пригнічення будуть успішними в довгостроковій перспективі; раніше не було ніяких спроб втручання державної системи охорони здоров'я з таким руйнівним впливом на суспільство протягом такого тривалого часу. Як населення і суспільство будуть реагувати на це, залишається неясним.

В даній дипломній роботі використовувались прості математичні моделі для оцінки можливих масштабів ураження та прогнозування характеристик пандемії, викликаной COVID-19.

Подальші дослідження повинні бути зосереджені на оновленні прогнозів з використанням нових даних і більш складних математичних моделей. Також подальших дослідженнях доцільно розширити номенклатуру моделей, які будуються безпосередньо на основі статистичних даних, зокрема моделей на основі методів штучного інтелекту. Також можна перевірити за допомогою імітаційного моделювання який процес поширення при взаємодії декількох популяцій (суспільств), як на швидкість поширення вплине зменшення контакту між популяціями і наскільки важливо вчасно обмежити цей контакт. Провести дослідження, як буде поширюватися вірус серед декількох великих міст з транспортними вузлами, які фактори зможуть вплинути або стати вирішальними проти пандемії.

ЛІТЕРАТУРА

1. Бідюк П.І., Романенко В.Д., Тимощук О.Л. Аналіз часових рядів. Київ: НТУУ «КПІ», Політехніка, 2013. 600 с.
2. Лукашин Ю. П. Адаптивні методи короткострокового прогнозування часових рядів [Текст] : навч. посіб. / Москва: Фінанси і статистика, 2003. 416 с.
3. Бокс Дж., Дженкінс Г. Анализ временных рядов. Москва: Мир, 1974. 406 с.
4. Бідюк П. І. Методи прогнозування. Том 1 і 2 / Луганськ : Альма Матер, 2008. 725с.
5. Бідюк П. І. Часові ряди: моделювання і прогнозування: монографія / Київ: ЕКМО, 2003. 144 с.
6. Бідюк П.І Аналіз часових рядів: навч. посіб. / Київ: НТУУ «КПІ», 2013. 600 с.
7. Holt C.C. Forecasting trends and seasonals by exponentially weighted moving averages / O.N.R. Memorandum, Carnegie Inst. of Technology. 1957.
8. Nesteruk, I. “Characteristics of coronavirus epidemic in mainland China estimated with the use of official data available after February 12, 2020.” [Preprint.] *ResearchGate*. 2020 Mar. doi:10.13140/RG.2.2.19667.32804.
9. Nesteruk, I. “Estimations of the coronavirus epidemic dynamics in South Korea with the use of SIR model” [Preprint.] *ResearchGate*. 2020 Mar. doi: 10.13140/RG.2.2.15489.40807.
10. Nesteruk, I. “Comparison of the coronavirus epidemic dynamics in Italy and mainland China” [Preprint.] *ResearchGate*. 2020 March. doi:10.13140/RG.2.2.19152.87049.
11. Атамась Н., Маєвська К. Вони дочекалися, або чому COVID-19 перший, але не останній, [електронний ресурс] – Режим доступу: <https://kunsht.com.ua/voni-dochekalisya/>

12. Георгій Ерман. BBC News Україна. Чума, холера, "іспанка": як великі пандемії змінювали світ. [Електронний ресурс] – Режим доступу: <https://www.bbc.com/ukrainian/features->
13. Щегольов І. Страшные эпидемии, укрощенные человеком. [Електронний ресурс] – Режим доступу: <https://rg.ru/2015/01/02/pandemia-site.html>
14. Hayes, J.N. Epidemics and Pandemics: Their Impacts on Human History [Текст] : 2005. 214-219 ст.

ДОДАТОК А (Статистичні дані)

24.01.2020 - 02.06.20	США	
1	1 827 688	463 868
2	1 806 950	458 231
3	1 766 253	416 461
4	1 766 253	416 461
5	1 742 178	406 446
6	1 718 074	399 991
7	1 695 587	391 508
8	1 677 400	384 902
9	1 658 854	379 157
10	1 639 869	366 736
11	1 619 325	361 239
12	1 597 720	350 135
13	1 574 047	298 418
14	1 548 812	294 312
15	1 525 588	289 392
16	1 505 423	283 178
17	1 483 936	272 265
18	1 465 056	268 376
19	1 440 109	250 747
20	1 415 176	246 414
21	1 387 906	243 430
22	1 367 104	232 733
23	1 345 455	230 287
24	1 326 895	216 169
25	1 307 210	212 534
26	1 281 607	198 993
27	1 254 828	195 036
28	1 226 471	189 910
29	1 202 268	189 791
30	1 178 373	187 180
31	1 156 073	180 152
32	1 130 623	175 382
33	1 101 727	164 015
34	1 067 726	153 947
35	1 038 321	120 720
36	1 011 027	115 936
37	986 653	111 424
38	964 257	106 988
39	936 692	100 372

26.02.20 - 02.06.20 Бразилія	Бразилія	
1	555 383	223 638
2	526 447	211 080
3	498 440	200 892
4	498 440	200 892
5	465 166	189 476
6	438 238	177 604
7	411 821	166 647
8	391 222	158 593
9	374 898	153 833
10	363 211	149 911
11	347 398	142 587
12	330 890	135 430
13	310 087	125 960
14	291 579	116 683
15	271 885	106 794
16	255 368	100 459
17	241 080	94 122
18	233 511	89 672
19	220 291	84 970
20	203 165	79 479
21	190 137	78 424
22	178 214	72 597
23	169 594	67 384
24	162 699	64 957
25	156 061	61 685
26	146 894	59 297
27	135 773	55 350
28	126 611	51 370
29	115 455	48 221
30	108 620	45 815
31	101 826	42 991
32	97 100	40 937
33	92 202	38 039
34	87 187	35 935
35	79 685	34 132
36	73 235	32 544
37	67 446	31 142
38	63 100	30 152
39	59 324	29 160

40	903 902	99 079
41	867 601	80 203
42	838 818	77 366
43	810 417	75 204
44	782 924	72 329
45	757 723	70 337
46	730 929	64 840
47	698 489	58 545
48	666 610	54 703
49	635 228	52 096
50	606 603	47 763
51	579 572	43 482
52	554 272	32 988
53	525 464	31 270
54	495 669	28 790
55	460 615	25 410
56	428 300	23 559
57	395 521	21 763
58	366 033	19 581
59	336 477	17 448
60	308 300	14 652
61	275 178	9707
62	243 195	9001
63	213 003	8474
64	187 862	7024
65	161 599	5644
66	140 726	2665
67	121 310	1072
68	101 527	869
69	83 727	681
70	65 690	361
71	53 665	348
72	43 604	171
73	33 795	171
74	25 478	171
75	19 073	147
76	13 663	108
77	7776	106
78	6413	17
79	4624	17
80	3491	12
81	2723	12
82	2179	12
83	1663	12

40	54 043	27 655
41	50 036	26 573
42	45 757	25 318
43	43 079	22 991
44	40 743	22 130
45	38 654	22 130
46	36 658	14 026
47	33 682	14 026
48	30 425	14 026
49	28 320	14 026
50	25 262	3046
51	23 430	173
52	22 192	173
53	20 727	173
54	19 638	173
55	18 092	173
56	16 170	127
57	14 034	127
58	12 161	127
59	11 130	127
60	10 360	127
61	9056	127
62	8044	127
63	6836	127
64	5717	127
65	4579	120
66	4256	6
67	3904	6
68	3417	6
69	2985	6
70	2554	2
71	2247	2
72	1924	2
73	1546	2
74	1021	2
75	793	2
76	621	2
77	372	2
78	321	2
79	200	1
80	162	
81	151	
82	151	
83	52	

84	1281	8
85	959	8
86	605	8
87	537	8
88	417	8
89	278	8
90	221	8
91	153	8
92	122	8
93	101	7
94	76	7
95	70	7
96	62	7
97	60	6
98	59	6
99	53	6
100	53	5
101	35	5
102	35	5
103	35	5
104	15	3
105	15	3
106	15	3
107	15	3
108	15	3
109	15	3
110	15	3
111	15	3
112	13	3
113	13	3
114	12	3
115	12	3
116	12	
117	12	
118	12	
119	12	
120	11	
121	11	
122	8	
123	8	
124	6	
125	5	
126	5	
127	5	

84	38	
85	31	
86	25	
87	20	
88	13	
89	13	
90	4	
91	4	
92	2	
93	2	
94	2	
95	2	
96	1	
97	1	
98	1	

128	5	
129	5	
130	2	
131	2	

01.02.20 - 02.06.20	Іспанія	
1	240 008	150 376
2	239 719	150 376
3	239 311	150 376
4	239 311	150 376
5	238 690	150 376
6	238 099	150 376
7	236 601	150 376
8	236 601	150 376
9	236 032	150 376
10	235 772	150 376
11	235 290	150 376
12	234 824	150 376
13	233 037	150 376
14	232 555	150 376
15	232 037	150 376
16	231 606	150 376
17	230 698	146 446
18	230 698	146 446
19	230 183	144 783
20	229 540	143 374
21	228 691	140 823
22	228 030	138 980
23	227 436	137 139
24	224 350	136 166
25	223 578	133 952
26	222 857	131 148
27	221 447	128 511
28	220 325	126 002
29	219 329	123 486
30	218 011	121 343
31	217 466	118 902
32	216 582	117 248
33	213 435	112 050
34	213 435	112 050
35	213 435	112 050
36	213 435	112 050
37	213 435	112 050

22.01.20 - 02.06.20	Китай	
1	83 023	78 317
2	83 022	78 316
3	83 001	78 305
4	83 001	78 305
5	82 999	78 302
6	82 995	78 291
7	82 995	78 288
8	82 993	78 280
9	82 992	78 277
10	82 985	78 268
11	82 974	78 261
12	82 971	78 258
13	82 963	78 241
14	82 963	78 241
15	82 963	78 241
16	82 963	78 241
17	82 954	78 238
18	82 947	78 227
19	82 941	78 219
20	82 933	78 209
21	82 929	78 195
22	82 926	78 189
23	82 919	78 171
24	82 918	78 144
25	82 901	78 120
26	82 887	77 993
27	82 886	77 993
28	82 885	77 957
29	82 883	77 911
30	82 881	77 853
31	82 880	77 766
32	82 875	77 685
33	82 875	77 679
34	82 874	77 642
35	82 862	77 610
36	82 858	77 578
37	82 836	77 555

38	213 435	112 050
39	213 435	95 708
40	213 435	92 355
41	213 024	89 250
42	208 389	85 915
43	204 178	82 514
44	200 210	80 587
45	198 674	77 357
46	191 726	74 797
47	190 839	74 797
48	184 948	74 797
49	177 644	70 853
50	172 541	67 504
51	170 099	64 727
52	166 831	62 391
53	163 027	59 109
54	158 273	55 668
55	153 222	52 165
56	148 220	48 021
57	141 942	43 208
58	136 675	40 437
59	131 646	38 080
60	126 168	34 219
61	119 199	30 513
62	112 065	26 743
63	104 118	22 647
64	95 923	19 259
65	87 956	16 780
66	80 110	14 709
67	73 235	12 285
68	65 719	9357
69	57 786	7015
70	49 515	5367
71	39 885	3794
72	35 136	3355
73	28 603	2125
74	25 374	2125
75	20 410	1588
76	17 963	1107
77	13 910	1081
78	11 748	1028
79	9942	530
80	7798	517
81	6391	517

38	82 830	77 474
39	82 827	77 394
40	82 819	77 357
41	82 804	77 257
42	82 790	77 157
43	82 779	77 125
44	82 747	77 093
45	82 735	77 068
46	82 718	77 029
47	82 694	77 003
48	82 341	76 936
49	82 294	76 905
50	82 249	76 861
51	82 159	76 799
52	82 085	76 773
53	81 969	76 738
54	81 907	76 707
55	81 865	76 660
56	81 804	76 622
57	81 739	76 582
58	81 707	76 541
59	81 668	76 517
60	81 638	76 472
61	81 623	76 436
62	81 589	76 408
63	81 555	76 242
64	81 524	76 062
65	81 478	75 784
66	81 444	75 454
67	81 401	74 972
68	81 345	74 594
69	81 298	74 055
70	81 221	73 655
71	81 180	73 163
72	81 116	72 709
73	81 060	72 252
74	81 014	71 749
75	80 977	71 158
76	80 931	70 430
77	80 906	69 653
78	80 884	68 700
79	80 867	67 816
80	80 848	66 926
81	80 827	65 572

82	5232	193
83	2277	183
84	2277	183
85	1695	32
86	1073	32
87	673	30
88	500	30
89	400	2
90	259	2
91	222	2
92	165	2
93	120	2
94	84	2
95	45	2
96	32	2
97	15	2
98	13	2
99	6	2
100	2	2
101	2	2
102	2	2
103	2	2
104	2	2
105	2	2
106	2	2
107	2	2
108	2	2
109	2	2
110	2	
111	2	
112	2	
113	2	
114	2	
115	2	
116	1	
117	1	
118	1	
119	1	
120	1	
121	1	
122	1	
123	1	

82	80 801	64 109
83	80 793	62 824
84	80 785	61 569
85	80 757	60 106
86	80 735	58 735
87	80 699	57 320
88	80 652	55 478
89	80 573	53 888
90	80 422	52 240
91	80 271	49 955
92	80 151	47 404
93	80 026	44 810
94	79 826	42 118
95	79 251	39 279
96	78 824	36 291
97	78 498	32 898
98	78 065	30 053
99	77 660	27 650
100	77 152	24 990
101	76 938	23 170
102	76 922	22 687
103	75 472	18 693
104	74 999	18 002
105	74 546	15 952
106	74 139	14 199
107	72 364	12 455
108	70 446	10 748
109	68 347	9294
110	66 292	7973
111	59 832	6213
112	44 699	5079
113	44 327	4635
114	42 306	3917
115	39 790	3218
116	36 778	2595
117	34 075	1998
118	30 553	1476
119	27 409	1115
120	23 680	843
121	19 693	614
122	16 607	463
123	11 871	275
124	9783	214
125	8124	135

126	6070	120
127	5494	101
128	2863	58
129	2062	49
130	1399	39
131	916	36
132	639	30
133	547	28

ДОДАТОК В (Презентація)

МОДЕЛІ ТА ПРОГНОЗИ ПРОЦЕСУ ПОШИРЕННЯ КОРОНАВІРУСУ

Виконала студентка IV курсу, групи КА-64, Мельничук Анастасія Володимирівна
Науковий керівник: професор, д.т.н., Бідюк П.І.

ТЕМА: МОДЕЛІ ТА ПРОГНОЗИ ПОШИРЕННЯ ПРОЦЕСІВ
ПОШИРЕННЯ КОРОНАВІРУСУ.

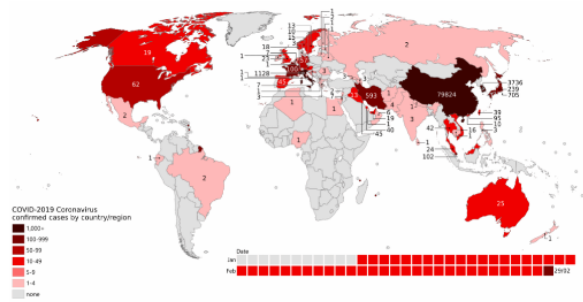
ОБ'ЄКТ ДОСЛІДЖЕННЯ: ПРОЦЕСИ РОЗВИТКУ ТА ПОШИРЕННЯ
ЕПІДЕМІЙ, ЯКІ БУЛИ СПРИЧИНЕНІ ВІРУСНИМИ
ЗАХВОРЮВАННЯМИ.

ПРЕДМЕТ ДОСЛІДЖЕННЯ: РЕГРЕСІЙНИЙ АНАЛІЗ
СТАТИСТИЧНИХ ДАНИХ, ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ ПРОЦЕСУ
ПОШИРЕННЯ ВІРУСУ В РІЗНИХ УМОВАХ.

МЕТА РОБОТИ: ОЦІНКА НАЯВНИХ СТАТИСТИЧНИХ ДАНИХ ТА
ПОБУДОВА РЕГРЕСІЙНИХ МОДЕЛЕЙ ДЛЯ ДАНИХ ЧАСОВИХ
РЯДІВ; ПОБУДОВА КОРОТКОЧАСНОГО ОДНОКРОКОВОГО
ПРОГНОЗУ КУМУЛЯТИВНОЇ КІЛЬКОСТІ ТИХ, ХТО ІНФІКОВАНИЙ І
ТИХ, ХТО ОДУЖАВ; ОЦІНКА ТА АНАЛІЗ ОСНОВНИХ МЕТОДІВ
БОРЬОБИ ПРОТИ ПОШИРЕННЯ ЕПІДЕМІЙ ВІРУСНИХ
ЗАХВОРЮВАНЬ.

ПОСТАНОВКА ЗАДАЧІ

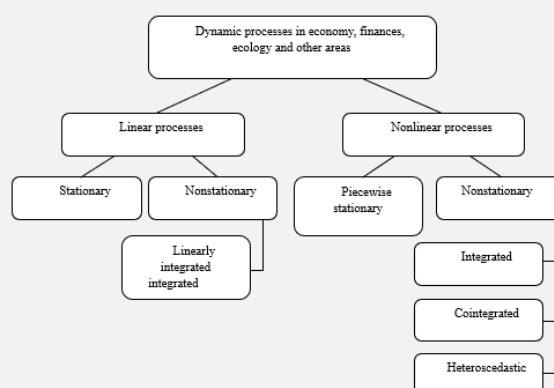
1. Знайти і підготувати статистичні дані стосовно кількості інфікованих та тих, хто вже одужав в Україні, Китаї, Іспанії, Бразилії, США. За необхідності виконати їх нормування чи логарифмування.
2. Попередньо проаналізувати знайдені часові ряди на наявність тренду.
3. Побудувати апроксимацію процесів поширення коронавірусу за допомогою моделей авторегресії та авторегресії з ковзним середнім.
4. Виконати оцінку короткострокових прогнозів процесів поширення коронавірусу.
5. Проаналізувати отримані результати, продемонструвавши їх на графіках.
6. Провести загальний аналіз умов розвитку та поширення інфекційного захворювання в популяції (в суспільстві).
7. Визначити основні фактори, які впливатимуть на швидкість розповсюдження вірусу та масштаби ураження, на основі яких сформулювати гіпотези.
8. Побудувати різного роду симуляції за допомогою модулів, на основі SIR-моделі, в програмі AnyLogic для підтвердження або спростування висунутих гіпотез та оцінювання ефективності основних методів боротьби проти поширення інфекційних захворювань.
9. Продемонструвати наглядно отримані результати, провести їх аналіз.



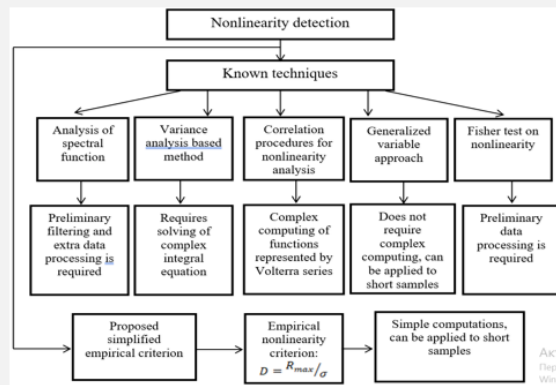
АКТУАЛЬНІСТЬ ПРОБЛЕМИ, ХАРАКТЕРИСТИКА РОЗВИТКУ ПРОЦЕСІВ ПОШИРЕННЯ ПАНДЕМІЙ В УКРАЇНІ ТА СВІТІ

ЧИ БУЛИ ПЕРЕДУМОВИ ДЛЯ ВИНИКНЕННЯ COVID-19 І
ЯКІ? ЧОМУ ВІН, НАЙІМОВІРНІШЕ, НЕ ОСТАННІЙ?

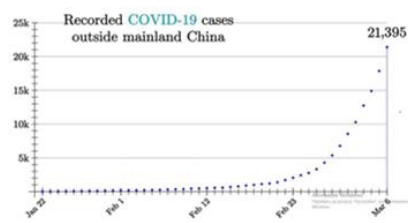
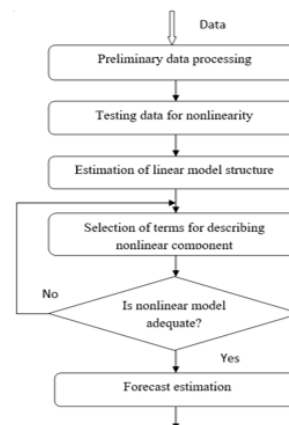
СПРОЩЕНА КЛАСИФІКАЦІЯ ДИНАМІЧНИХ ПРОЦЕСІВ



ДЕЯКІ МЕТОДИ ТЕСТУВАННЯ ДАНИХ НА НЕЛІНІЙНІСТЬ



ПРОЦЕДУРА, ЯКА
ІЛЮСТРУЄ ПОШУК
ФОРМАЛЬНОГО
ОПISУ НЕЛІНІЙНОГО
ПРОЦЕСУ



$$S(t) + I(t) + R(t) = \text{Constant} = N$$

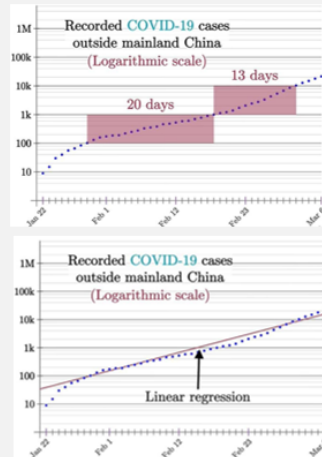
SIR-МОДЕЛЬ

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0,$$

$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$

$$\frac{dR}{dt} = \gamma I.$$

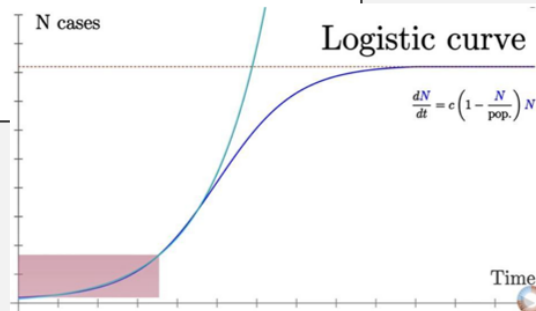


- 30 днів за підрахунками по нашій моделі число інфікованих дорівнює 1 млн.
- (насправді станом на 5 квітня світова статистика говорила про 1 202 827 осіб, з яких на той момент уже померло 64 774).
- 47 днів це число має досягти відмітки 10 млн.
- (в реальності станом на 22 квітня число заражених по світу склало всього 2 567 495 осіб, з яких тоді померло 177 935)
- 64 дні отримуюмо результат рівний 100 млн. інфікованих.
- (реальна картина на 9 травня склала 4 023 969 інфікованих та 279 311 померло)
- 81 день і число заражених мало б досягнути відмітки в 1 млрд. осіб.
- (насправді станом на 26 травня 5 590 364 зареєстрованих випадків і 352 093 смерті)

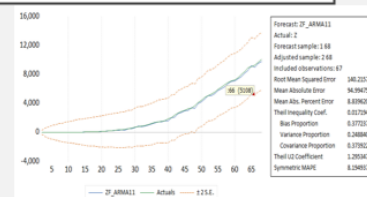
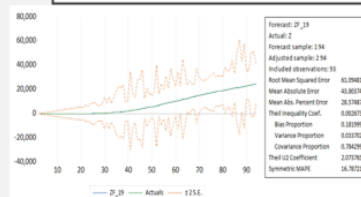
РОЗГЛЯНЕМО МОЖЛИВІ ДВІ ФУНДАМЕНТАЛЬНІ СТРАТЕГІЇ БОРОТЬБИ З ПОШИРЕННЯМ ПАНДЕМІЇ КОРОНАВІРУСА:

- A) ПРИГНІЧЕННЯ.
B) ПОМ'ЯКШЕННЯ.

$$\text{Growth factor} = \frac{\Delta N_d}{\Delta N_{d-1}}$$



КОРОНА-ВІРУС В УКРАЇНІ ЗА ПЕРІОД З 23.03.2020 ПО 02.06.2020



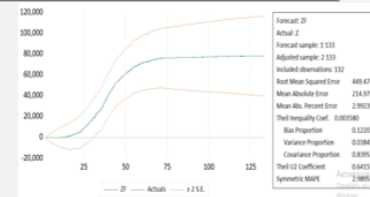
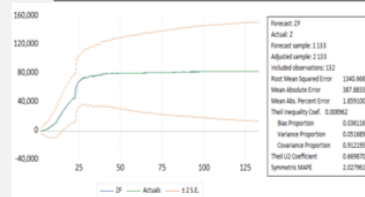
ХАРАКТЕРИСТИКА НАЙКРАЩОЇ ЗА ЯКІСТЬ ПРОГНОЗУ КІЛЬКОСТІ ЗАХВОРИВШИХ МОДЕЛІ АРКС(1,9), DW=1,842

$$y(k) = 12329,27 + 0,998y(k-1) + mv(k) + 1,1277mv(k-1) + 1,3501mv(k-2) + 1,1226mv(k-3) + 1,1911mv(k-4) + 1,1949mv(k-5) + 1,1786mv(k-6) + 1,3712mv(k-7) + 1,0131mv(k-8) + 0,2683mv(k-9)$$

ХАРАКТЕРИСТИКА НАЙКРАЩОЇ ЗА ЯКІСТЬ ПРОГНОЗУ КІЛЬКОСТІ ОДУЖАВШИХ МОДЕЛІ АРКС(1,1), DW=0,999

$$y(k) = 5045,23 + 0,999y(k-1) + mv(k) + 0,748mv(k-1)$$

КОРОНА-ВІРУС В КИТАЇ ЗА ПЕРІОД З 22.01.2020 ПО 02.06.2020



ХАРАКТЕРИСТИКА НАЙКРАЩОЇ ЗА ЯКІСТЬ ПРОГНОЗУ КІЛЬКОСТІ ЗАХВОРИВШИХ МОДЕЛІ $APKC(1,10)+KC(15)$, $DW=1,9995$

$$y(k) = 49451,33 + 0,995y(k-1) + mv(k) + 0,443mv(k-1) + 0,199mv(k-2) + 0,241mv(k-3) + 0,307mv(k-4) + 0,234mv(k-5) + 0,175mv(k-6) + 0,111mv(k-7) + 0,107mv(k-8) + 0,224mv(k-9) + 0,003mv(k-10) - 0,104mv(k-15)$$

ХАРАКТЕРИСТИКА НАЙКРАЩОЇ ЗА ЯКІСТЬ ПРОГНОЗУ КІЛЬКОСТІ ОДУЖАВШИХ МОДЕЛІ $APKC(1,5)$, $DW=1,778$

$$y(k) = 40092,21 + 0,999y(k-1) + mv(k) + 0,474mv(k-1) + 0,868mv(k-2) + 0,738mv(k-3) + 0,446mv(k-4) + 0,354mv(k-5)$$

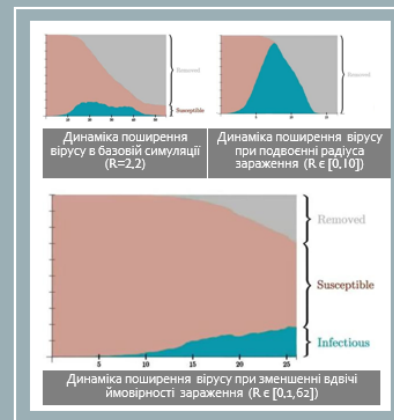
SIR-МОДЕЛЬ

В SIR-модель передбачено поділ популяції на три категорії:

- S-susceptible;
- I-infectious;
- R-recovered.

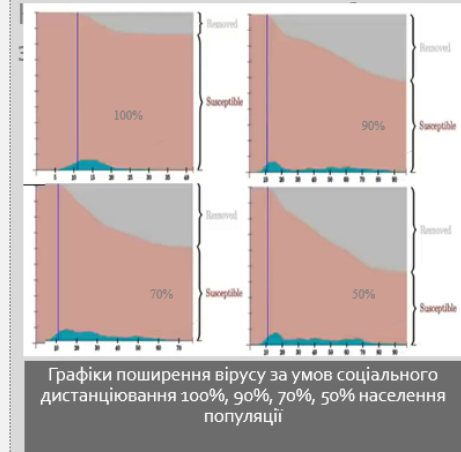
Створено базову симуляцію, яка своїми показниками наближено нагадує динаміку поширення коронавірусу.

Виконано перевірку на логіку, результати зображенні на графіку:



МОДЕЛЮВАННЯ СИМУЛЯЦІЇ З ЦЕНТРАЛЬНОЮ ТОЧКОЮ ВІДВІДУВАННЯ

ВПЛИВ СОЦІАЛЬНОГО
ДИСТАНЦІЮВАННЯ НА
ДИНАМІКУ ПОШИРЕННЯ ВІРУСУ
СЕРЕД НАСЕЛЕННЯ ПОПУЛЯЦІЇ



ДЯКУЮ ЗА УВАГУ!